SCIENCE LIBRARY 18 OCT 5 - 1931

Vol. XIV, No. 5 PSYCHOLOGICAL REVIEW PUBLICATIONS OCTOBER, 1931

Journal of Experimental Psychology

EDITED BY

SAMUEL W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Monographs*)
EDWARD S. ROBINSON, YALE UNIVERSITY (*Bulletin*)

CONTENTS

A Study of Certain Features of Punishment in Serial Learning: W. McTEER 453

Letter-Position Effects, in the Range of Attention Experiment, as Affected by the Number of Letters in Each Exposure: H. R. CROSLAND... 477

A Further Study of Configurational Learning in the Goldfish: F. T. PERKINS 508

The Memory Value of Size: T. KARWOSKI..... 539

The Aptitude Hypothesis in Motor Skills: S. SEASHORE 555

Group Judgments in the Field of Personality Traits: M. SMITH 562

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter, July 1, 1920, at the post-office at Lancaster, Pa.

PUBLICATIONS

OF THE

AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
S. W. FERNBERGER, UNIVERSITY OF PENNSYLVANIA (*J. Exper. Psych.*)
WALTER S. HUNTER, CLARK UNIVERSITY (*Index and Abstracts*)
HENRY T. MOORE, SKIDMORE COLLEGE (*J. Abn. and Soc. Psychol.*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY (*Monographs*)
EDWARD S. ROBINSON, YALE UNIVERSITY (*Bulletin*)

HERBERT S. LANGFELD, Business Editor

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bi-monthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 540 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly (10 numbers), the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bi-monthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 600 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL ABSTRACTS

appears monthly, the twelve numbers and an index supplement making a volume of about 700 pages. The journal is devoted to the publication of non-critical abstracts of the world's literature in psychology and closely related subjects.

PSYCHOLOGICAL MONOGRAPHS

consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY

appears quarterly, April, July, October, January, the four numbers comprising a volume of 448 pages. The journal contains original contributions in the field of abnormal and social psychology, reviews, notes and news.

ANNUAL SUBSCRIPTION RATES

Review: $5.50 (Foreign, $5.75). **Index:** $4.00 per volume.
Journal: $6.00 (Foreign, $6.25). **Monographs:** $6.00 per volume (Foreign, $6.30).
Bulletin: $6.00 (Foreign, $6.25). **Abstracts:** $6.00 (Foreign, $6.25).
Abnormal and Social: $5.00 (Foreign, $5.25). Single copies $1.50.
Current numbers: **Review or Journal, $1.00; Abstracts, 75c; Bulletin, 60c.**

COMBINATION RATES

Review and Bulletin: $10.00 (Foreign, $10.50).
Review and J. Exp.: $10.00 (Foreign, $10.50).
Bulletin and J. Exp.: $11.00 (Foreign, $11.50).
Review, Bulletin, and J. Exp.: $15.00 (Foreign, $15.75).
Review, Bulletin, J. Exp., and Index: $18.00 (Foreign, $18.75).

Subscriptions, orders, and business communications should be sent to the

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, N. J.

Journal of
Experimental Psychology

VOL. XIV, No. 5 OCTOBER, 1931

A STUDY OF CERTAIN FEATURES OF PUNISHMENT IN SERIAL LEARNING

BY WILSON McTEER[1]

College of the City of Detroit

CHAPTER I. INTRODUCTION

Watson (**8**) has advanced the theory that 'conditioning' to the 'social don'ts' would be accomplished best by punishment direct from the perceptual situation. There are at least six independent variables to be considered in making an experimental test of this theory: (*a*) The subject, who may vary in sex, age, experience, intelligence, etc; (*b*) The activity pattern desired, whether acquisitive or eliminative, slow or fast, etc; (*c*) The punishment used, whether intense or mild, whether automatic or administered by some person, etc; (*d*) The spatial relation of the punishment to the reacting member; (*e*) The temporal relation of the punishment to the act; (*f*) The subject's perceptual understanding of the relation of the punishment to the act.

Since the learning problem used in this experiment is a new one, the principal question investigated in this research is to determine whether electric shock punishment in a ten-unit finger-response multiple-choice serial-order learning problem would result in more or less efficient learning than when no punishment is used.

Secondarily, the present experiment is concerned with the

[1] From the Psychological Laboratory of the University of Chicago. The writer wishes to express his appreciation to Professor A. G. Bills for suggesting the problem and for his valuable counsel during the course of the experimentation.

fourth variable mentioned above. The attempt is made to determine whether punishment applied to the ankle in this type of learning would result in more or less efficient learning than punishment applied to the reacting finger.

Several investigators have dealt with the more general problem of electric shock punishment in various types of situations. Johanson (**3**) found a saving of 15 percent in reaction time when automatic punishment for long reactions was introduced. Warner Brown and collaborators (**1**) standardized Hamilton's 'multiple choice problem for human subjects' as outlined by Yerkes (**10**) and added electric shock for error. They report (**9**) that with the electric shock the problem becomes quite a sensitive measuring instrument, but do not give data comparing punished and non-punished groups.

Rexroad (**4, 5**) used punishment in a problem in which the five fingers of one hand were each to be associated with a different color and the correct finger was to press a key each time the appropriate color was flashed on a screen. An automatic record was taken of the correct responses and errors that occurred within a constant period of time. Three problems were studied. (*a*) In a study of the initial adjustment to the situation the punished group made on the average 2 percent more correct responses and 16 percent fewer errors than the control group. (*b*) A study of a small practiced group using the same combination at each sitting showed a decrease of 2 percent in correct responses and of 15 percent in errors for the days when punishment was used as compared with the days when no punishment was used. (*c*) A study of another small practiced group which learned a new color-finger-response combination at each sitting showed a decrease of 11 percent in correct responses and of 4 percent in errors for the punishment days as compared with the non-punishment days. In general, one may say that the effect of the punishment in all conditions was to make the subjects more cautious and slower in reacting.

Vaughn (**6**) investigated the effect of different types of instruction upon quickness of stopping an arm movement in

response to a 'danger signal' or red light. Electric shock punishment for slowness in stopping was the most effective form of 'instruction'. Next in order came (*a*) threat of punishment with occasional shock; (*b*) threat of punishment with a demonstration of how it might occur; (*c*) indefinite threat that something might happen; (*d*) simple 'don't' instruction; (*e*) positive instructions were least effective of all the forms used. He, however, called particular attention to the great variation between individuals and their reaction to various forms of instruction.

Bunch (**2**) introduced punishment for errors into maze learning. Punishment occurred only if the subject went to the end of the blind alley. He found that the punished group learned with fewer errors, fewer trials and less time. He also found that the most effective learning occurred with those subjects who reported that the experience of the shock was only moderately intense.

Vaughn and Diserens (**7**) have made a special study upon the effect of varying the intensity of the punishment in learning a simple maze. Their maze was constructed so that punishment was received in all parts of the blind alleys. Four patterns were learned under the following conditions of punishment: none, slight, moderate, intense. Each subject served in each condition, and a counterbalancing technique was used so that practice and maze differences supposedly 'washed out' in the final averages. They found fewest trials, fewest errors, and least time with *slight* punishment. Their apparatus was so constructed that they could determine the reaction time of the subject's response to the blind alley, and they report the quickest reaction time with the most intense punishment. They call particular attention to the marked individual differences in reaction to punishment.

Two unpublished studies were carried out at Wisconsin under Hull in '26, '27, and '28 which bordered upon the field of the secondary problem of the present investigation. The problem studied was the effect of punishment as applied to the reacting hand in maze learning compared with punishment applied to the non-reacting hand. The first investigator

claimed to find significant differences in favor of punishing the reacting hand. The second investigator repeated the experiment under more carefully controlled conditions and found practically no differences between the groups she used.

In general, it has been found that punishment of moderate intensity results in more efficient responding in reaction time, in maze learning, and in finger-response-multiple-choice of the types used by both Yerkes and Rexroad. Punishment of severe intensity is reported as being disruptive. The studies of the spatial relation of the punishment to the reacting member report conflicting results and consequently the question remains unsettled.

Chapter II. Procedure and Subjects

The present experiment was originally outlined with the view of attempting to determine the effect of punishment in the serial order learning of non-sense syllables. Actual experimentation with the apparatus which was designed for this purpose, however, resulted in the substitution of number combinations in preference to non-sense syllables.

The loci of punishment were selected with the view, first, of choosing spots of distinct anatomical and experiential separation; secondly, of selecting areas such that the electrodes could be quickly and conveniently attached without undue embarrassment to naive subjects. The finger and the ankle were the regions best suited to these requirements.

Extended experimentation of more than a year's duration was involved in perfecting the mechanical apparatus, in selecting the electrodes (particularly the ankle contact), in attempting to provide a reasonably constant induction shock punishment; in devising a scale for identifying as nearly as possible the subject's experience of the punishment; in selecting an objectively constant variation in the current that would result in practical equality of the subjective reports of the intensity of the experience; and, in simplifying the learning procedure involved so that all the subjects would follow approximately the same plan in learning, with the exception of those differences resulting from the punishment and from individual eccentricities.

The procedure as finally adopted was as follows: The subject upon entering the room was handed a sheet of typewritten instructions (see Appendix A) which, (*a*) emphasized learning with few repetitions; (*b*) pointed out that a sound from the box before him was the signal of the choice of the correct key; (*c*) suggested that a grouping of the key numbers as found would result in quickest learning; (*d*) warned him that there might be electric shock punishment for errors; (*e*) asked him to pause upon completing a trial until the experimenter signalled him to continue; (*f*) told him that the criterion of complete learning was two errorless trials; and (*g*) requested him not to discuss the experiment with others.

After the subject had read these instructions he was permitted to ask questions concerning any of the details that were not clear. He was then seated at the end of the table and the electrodes were fastened on his right ankle, calf, and wrist regardless of the experimental group to which he belonged.[2]

The experimenter then seated himself at the side of the table, adjusted the induction coil and the punishment circuit to the settings appropriate to the experimental group to which the subject belonged; placed the headphone 'error indicator' over his left ear, set the stopwatches and signalled the subject to start.

The problem the subject faced may best be thought of in terms of a typewriter. Assume a typewriter with only ten keys. Assume further that the typewriter is so built that the carriage will only advance when the right letter of the right word is struck. Suppose the word that this ten-letter machine is set for is 'republican'. Then, the subject's problem is to find the word by the simple process of pounding the keys. Suppose the keys are labeled in the order 'a, b, c, e, i, l, n, p, r, and u'. The subject in starting the first trial, then, might hit the keys a, c, l, *r*, and at *r* the carriage would move, indicating that it was the first letter of the word; u, p, n, l, i, *e*, might come next, and the subject would remember *r*, *e*; the, possibly, i, l, n, *p*; and so through the list until the subject had

[2] See details in W. McTeer, A New Device for Serial Order Learning, *this Journal*, 1931, 14, 446–452.

found the complete correct order *r*, *e*, *p*, *u*, *b*, *l*, *i*, *c*, *a*, *n*. The letters struck out of order would count as errors. In the later repetitions or trials the subject would attempt to spell the word correctly each time, but would repeat it until he had done so twice in succession.

The problem the subject faced in this experiment differed from the above hypothetical machine in that (*a*) a ten-unit number was learned instead of a word, (*b*) the 'advancement' was indicated by the tap of a sound hammer rather than by the movement of a typewriter carriage, (*c*) the keyboard was perpendicular to, rather than parallel with, the plane of the subject's body, and (*d*) in two of the three groups the subject received punishment for errors.

The following combinations were selected and used as presenting problems of approximately equal difficulty: first day: 3 8 2, 10 5 4, 9 7 1 6; second day: 1 9 4, 5 7 10, 2 6 8 3; third day: 8 4 9, 1 3 7, 6 10 2 5.

The experimenter, in addition to keeping the records, watched the subject's hand to prevent an adjustment which seemed to occur regularly in the Finger group. These subjects had the tendency to bend the finger so that they were striking the keys with the finger-nail rather than with the fleshy lobe of the finger. This adjustment tended to reduce the intensity of the punishment and so the subjects would increase their speed of tapping the keys. When this happened the experimenter would warn the subject at the end of the trial that the use of the finger-nail was not permitted.

The following records were kept during the course of the learning: (*a*) time and errors per 'correct' key in the first trial; (*b*) errors per 'correct' key and total time per trial in the later trials; (*c*) relevant comments and unusual events were noted when the speed of the subject permitted. These records were taken each day on a special mimeographed form which also included the subject's name, school classification, date, hour, the symbol of the list learned, the experimental group to which the subject was allocated, and a scale for recording the subject's judgment as to the intensity of the punishment experienced.

The intensity scale presented to the subject was a separate typewritten sheet with a vertical rectangle suggesting a foot rule. Nine equal steps were marked off on this rectangle and labeled as follows: (0) no sensation; (1) barely feel the current; (2) slight tingling; (3) tingling; (4) strong tingling; (5) slight pain; (6) painful; (7) very painful; (8) extremely painful. The subject expressed his judgment in terms of the numerals.

149 subjects took part in the experiment. 34 served in preliminary experiments in perfecting the technique and apparatus, while the records from 25 subjects who participated in the main experiment could not be used. The reasons for discarding these latter records may be summarized as follows: 11 misinterpreted the general problem on the first day and failed to learn it until the second or third day; 5 were lost due to apparatus difficulties; 3 failed to return on the second or third day.

The data presented are based upon records taken on 90 male subjects, each of whom learned three combinations, one per day at the same hour on three successive days (slight variations in time of work occurred with a few subjects, but these are equally distributed between the experimental groups). These subjects varied in age from 17 to 35 with the average age probably around 24 years. They varied in scholastic standing from sophmore to 3rd year graduate students.

These 90 subjects were distributed equally between the three experimental groups: (1) Control, or no punishment; (2) Ankle group, or punishment on the ankle; and (3) Finger group, or punishment on the finger. The three groups were equated as to age and scholastic standing. Seasonal and daily variations that might have influenced the intensity of the punishment, and variations in the efficacy of the instruction that might have crept in during the year that these data were being collected were effectually counterbalanced by following the plan of distributing each set of subjects equally between the three experimental groups.

Chapter III. Data and Discussion

The present experiment, as stated in the Introduction, was organized to investigate two questions: first, to determine the effect of punishment upon efficiency of learning in a ten-unit finger-response multiple-choice serial-order learning problem; second, to determine whether punishment applied to the ankle would result in more or less efficient learning in this type of problem than when the punishment is applied to the reacting finger.

The data collected permit two general approaches to the study of the questions investigated: first, in terms of the general criteria of total trials, total time and total errors; second, in terms of the special features of this type of problem.

The first approach permits the investigation of the two principal questions with reference to the following subsidiary problems:

1. The variable effect of punishment upon the respective criteria of learning (*i.e.* whether it has the most effect upon trials, time or errors);

2. The effect of punishment upon adaptation to a novel problem;

3. The day to day adaptation to a problem involving punishment;

4. The relation of the intensity of punishment to the criteria of learning.

The second approach involves an analysis of the situation faced in learning each combination. Such an analysis indicates that each subject had two distinct adaptations to make: first, exploration and memorization; that is, in the first trial he had to continue to tap the keys until he had found each unit of the combination and had attempted to memorize it in its proper order; second, retention and correction; that is, in the later trials he had to attempt to retain the correct combination in the proper order, or, failing in this, he had to correct any 'errors in order' [3] with as few trials and as little exploration as possible. Obviously, punishment might be a

[3] The phrase 'errors in order' is used here to distinguish errors in memorization rom the errors made in the exploratory search for a correct key.

differentiating factor in these two parts of the problem. This analysis then suggests the further study of the two main questions in relation to:

5. The effect of punishment during the exploratory trials; and

6. The effect of the punishment during the later trials.

The statistical treatment of the data in the present experiment has involved: (*a*) The calculation of the means and the evaluation of the differences between the means in terms of the probable error of the difference. The probable errors of the mean were calculated by the use of th formuls, P.E. = .6745 $\sigma/\sqrt{n}$. (*b*) The calculation of the standard deviations and the evaluation of the differences in spread in terms of the probable error of the difference between the standard deviations. (*c*) Day to day consistency and trend have also been used in evaluating the data, since most of the differences obtained do not meet the present accepted statistical criterion of being three times their own probable errors.

Contrary to the practice usually followed in dealing with learning data, the time for each correct trial and the two correct trials have been included in their respective totals. This is justifiable since the time in the correct trial might possibly vary with the experimental condition or group.

TABLE I*A*

AVERAGE TIME, AVERAGE TRIALS AND AVERAGE ERRORS WITH STANDARD DEVIATIONS FOR EACH EXPERIMENTAL GROUP ON EACH DAY

	Day I		Day II		Day III	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Time, in seconds						
Ankle group.....	546.0	291.1	313.0	211.5	285.7	136.9
Finger group.....	522.4	358.5	300.4	121.3	263.2	124.7
Control group.....	546.9	347.6	272.2	125.2	302.0	175.5
Trials						
Ankle group.....	7.04	4.30	6.70	3.65	6.53	3.09
Finger group.....	8.60	4.98	7.30	3.22	6.83	3.16
Control group.....	9.96	5.13	7.93	4.34	7.96	4.45
Errors						
Ankle group.....	101.8	63.4	50.6	31.9	46.2	27.4
Finger group.....	127.7	122.5	55.8	34.8	46.2	29.4
Control group.....	129.8	81.9	74.6	71.8	69.9	46.4

TABLE I*B*

THE COEFFICIENTS OF THE RELIABILITY OF THE DIFFERENCES BETWEEN THE MEANS GIVEN IN TABLE I*A*

	Day I	Day II	Day III
Time			
Ankle-Finger	.41	.44	.99
Control-Finger	.39	−1.30	2.22
Control-Ankle	.016	−1.39	.59
Trials			
Finger-Ankle	1.80	1.00	.56
Control-Finger	1.54	.94	1.67
Control-Ankle	3.53	1.70	2.16
Errors			
Finger-Ankle	1.86	.90	.00
Control-Finger	.12	1.91	3.51
Contro Ankle	2.20	2.48	3.57

TABLE I*C*

THE COEFFICIENTS OF THE RELIABILITY OF THE DIFFERENCES BETWEEN THE STANDARD DEVIATIONS GIVEN IN TABLE I*A*

	Day I	Day II	Day III
Time			
Ankle-Finger	−1.68	4.20	.76
Control-Finger	−0.25	.25	2.74
Control-Ankle	1.43	−4.03	2.00
Trials			
Finger-Ankle	1.21	−1.02	.18
Control-Finger	.24	2.38	2.74
Control-Ankle	1.43	1.40	2.95
Errors			
Finger-Ankle	4.90	.70	.55
Control-Finger	−3.15	5.32	3.50
Control-Ankle	2.04	5.80	4.04

Section 1.—The variable effect of punishment upon the respective criteria of learning:

Three relatively independent criteria of learning efficiency were used in this experiment. It is possible that punishment for errors might produce any one of the following results: (*a*) all criteria might be equally affected; (*b*) two of the criteria might be more definitely affected than the third; (*c*) one of the criteria might be markedly affected while the others were only slightly disturbed. An examination of Tables I*A*, I*B* and I*C* above will show that this latter possibility is the truer statement of the present results. Error averages were more

definitely and consistently affected by the punishment than were the time averages or the trial averages.

This is in agreement with the findings of Rexroad, Bunch, and Vaughn in so far as it applies to the reduction of errors with punishment. Bunch, using maze learning, also found a marked reduction in time and trials.

Section 2.—The effect of punishment upon adaptation to a novel problem:

All subjects in all three groups read the same instructions. If the fear of the electric shock as suggested by the instructions rather than the shock itself were the dominant factor in the adjustment, the three groups should have been approximately equal in all criteria on the first day. But, a careful examination of the results obtained on the first day (Tables I*A*, I*B*, I*C*) will show that there are differences between the three groups. In average time there are no marked differences, but in average trials and average errors there are distinct differences. The punished groups require fewer trials to learn, and somewhat fewer errors.

A comparison of the means of the finger and the ankle groups shows that the ankle group required fewer trials to learn and made fewer errors than did the finger group. The experimenter's observation of the behavior adjustments made by the members of the two groups during the learning period on the first day indicates that the obtained differences are due to differences in the experimental situation rather than to a chance superiority of one group. It was noted very early in the course of the experiment that the members of the finger group very quickly adapted to the situation by tapping the keys quite hard and by withdrawing the hand very quickly; while the members of the ankle group, on the other hand, tended to press the keys and to hold them down a fraction of a second longer than did the members of the finger group. The effect of this difference in behavior adjustment was that certain of the finger-punished individuals tended to lose track of the key they had just touched, while the ankle-punished individuals, on the contrary, took time to observe the key touched in reference to the other keys but experienced a longer and more intense punishment.

Since the total time is approximately the same with all three groups and yet the punished groups made fewer errors, it is possible that the effect of the punishment was to produce more cautious behavior. An index of speed, or caution, is made possible by the nature of the experiment since a key has to be touched before either an advancement or an error may occur. The average number of seconds per response or per key touched may be obtained by the use of the formula:

$$\frac{\text{Total Time}}{\text{Total Errors} + 10\,(\text{Total Trials})}.$$

By the use of this formula the first day averages for each of the three groups were found to be: Ankle 3.26 ($\pm$.17) seconds per response; Finger 3.03 ($\pm$.15) seconds per response; and, Control 2.61 ($\pm$.19) seconds per response. The differences between the punished groups and the control group are about two times their probable errors.

In so far as punishment results in more cautious behavior, these findings verify the earlier experiments of Johanson, Rexroad, Bunch, Vaughn, and Vaughn and Diserens. Whether cautions behavior, however, produces beneficial or detrimental results probably depends upon the nature of the problem and the criteria used.

Section 3.—The day to day adaptation to a problem involving punishment:

Logically, there are three possibilities as to the part the punishment might play throughout the three days: (*a*) the groups might maintain their same relative position as on the first day; (*b*) they might become further separated; or (*c*) the punishment factor might be adapted to and the behavior become more similar. All three effects were obtained, but with different criteria and different groups.

Referring to the tables again, it may be observed that so far as time averages are concerned there are no consistent relations between the punished groups and the control group. However, a comparison of the means of the ankle and the finger groups shows a slight but consistent difference in favor of the finger group. Each group shows a marked reduction

in average time on the second and third day as compared with its own first day average.

The punished groups consistently learn in fewer trials and with less variability than the control group on all three days. It may also be noted that the finger group improves more rapidly, relatively speaking, than the ankle group. On the third day both punished groups learn the combination in approximately the same number of trials.

The punished groups show a marked day to day improvement over the control group in errors. The differences obtained on the third day are three times their probable errors. It may also be noted that the punished groups show consistently less variability in errors on the second and third days; here again the differences obtained are three, four and five times their probable errors. A more detailed description of the adjustments that resulted in these differences will be given in Sections 5 and 6.

The finger punished group, as noted in Section 2, made almost as many errors as did the control group on the first day. On the second day they made only slightly more errors on the average than did the ankle group. Both the finger and the ankle groups made the same average number of errors on the third day. Here, as in trials, punishment tends to produce greater uniformity of behavior.

Section 4.—The relation of the intensity of punishment to the criteria of learning:

Bunch, using a ten-point scale varying between the extremes of (1) 'least perceptible shock', and (10) 'most one could stand', found the most efficient learning with those subjects that reported an intensity of (4) and (5). His objective current was kept constant. In view of this report, scatter-diagrams were plotted with the data obtained in this experiment. There was so little evidence of relation in these diagrams that no attempt was made to calculate correlation coefficients. However, the present results are not to be considered as contradictory to the results obtained by Bunch, or by Vaughn and Diserens, since the scale used in estimating the punishment, the objective intensities used, the manner of

application, and the problem to be solved differed in each experiment.

TABLE II

AVERAGE TIME PER RESPONSE IN RELATION TO AVERAGE SUBJECTIVE INTENSITY

	Day I	Day II	Day III
Time per response			
Ankle group	3.26	2.72	2.74
Finger group	3.04	2.48	2.50
Control group	2.61	2.23	2.22
Intensity reported			
Ankle group	4.08	4.01	3.76
Finger group	3.80	3.48	3.33
Control group	0	0	0

Though there seems to be little correlation within either group between intensity of punishment and learning efficiency, it is interesting to note that the day to day differences between the finger and ankle groups in 'number of seconds per response' (see Section 2) are correlated with differences in the reported intensity of experience as shown in Table II above.

A careful study of this table suggests that the differences between the finger and the ankle groups which were discussed in Sections 2 and 3 might be primarily due to differences in intensity of punishment. To investigate this possibility a study was made in which reported intensity was held constant by selecting the records of those individuals in each group that could be paired with a member of the other group that reported a punishment experience of the same intensity. The results obtained are given in Table III.

TABLE III

AVERAGE SCORES WHEN REPORTED INTENSITY OF PUNISHMENT IS HELD CONSTANT

	Day I	Day II	Day III
Number in groups	24	21	21
Average intensity	4.04	3.5	3.7
Average time: A	539.9	286.8	266.3
F	533.6	280.5	228.0
Average trials: A	7.29	6.5	5.76
F	8.5	7.0	5.76
Average errors: A	91.6	49.6	41.3
F	108.7	52.0	38.1

In the above table it is interesting to note that the differences discussed in the preceding sections still remain after

punishment is equated, though they are greatly reduced in magnitude and statistical significance. This would suggest the conclusion that the finger-ankle spatial factor is relatively unimportant in this type of problem beyond the first day.

That in Table II the averages of the reported intensity should vary after the preliminary experiments in which objective values were selected that were experienced as subjectively equal, is probably due to two peculiarities of this experimental set-up: (*a*) As previously mentioned in Section 2 the ankle-punished individuals tended to hold down the keys longer than did the finger-punished individuals. (*b*) In addition, the subjective experience of the finger-punishment was not quite the same as that of the ankle-punishment. The experience of the finger-punishment may be described as a painful tingle and muscle-jerk which spread through the finger and back of the hand and persisted only while the current was passing. The ankle-punishment may better be described as an half-inch area of hot sharp needle-like pains which persisted for a brief time after the current ceased. Both groups experienced the electric shock as punishment and both considered it very unpleasant, but the persistence of the pain of the ankle-punishment tended to influence the subjects of that group to report it as being the more intense.

Section 5.—The effect of punishment during the exploratory trial:

The first trial in learning any of the combinations, as pointed out in the preliminary discussion of the results, differs markedly from the later trials in the nature of the adaptation which is required of the subject. Exploration is required, which may be cautious and logical or rapid and random, and memorization. If the exploration is cautious and logical, the subject will get through the combination with a minimum of errors, though probably with an increase in time. If the memorization is thorough, it is probable that the subject will make the two errorless runs very quickly after the initial trial. Thorough memorization, however, would also probably require additional time in the first trial.

In Tables IV *A* and IV *B* it is clearly evident that the

TABLE IV*A*

TIME AND ERROR AVERAGES MADE IN THE INITIAL TRIAL IN LEARNING EACH COMBINATION

	Day I	Day II	Day III
Time in seconds			
Ankle group	184.4 (±11.70)	116.0 (±5.08)	119.9 (±6.29)
Finger group	178.5 (±10.37)	109.2 (±4.94)	104.0 (±4.45)
Control group	150.4 (±13.60)	101.8 (±7.59)	92.4 (±6.91)
Errors			
Ankle group	45.7 (±2.53)	28.3 (±1.12)	28.1 (±1.14)
Finger group	43.8 (±2.25)	29.5 (±1.35)	29.3 (±1.34)
Control group	50.8 (±3.09)	36.0 (±3.07)	31.9 (±1.44)

TABLE IV*B*

THE COEFFICIENTS OF THE RELIABILITY OF THE DIFFERENCE BETWEEN THE MEANS GIVEN IN TABLE IV*A*

	Day I	Day II	Day III
Time			
Ankle-Finger	.30	.96	2.06
Finger-Control	1.64	.80	1.40
Ankle-Control	1.89	1.54	2.93
Errors			
Finger-Ankle	−.58	.70	.72
Control-Finger	1.85	1.62	1.29
Control-Ankle	1.27	2.34	2.07

punished groups required more time, but made fewer errors than the control group in the initial trial on each day. This would indicate that the addition of punishment to the experimental situation resulted in a more cautious approach on the part of the punished individuals.

A comparison of the finger and ankle groups in the same tables shows that the finger group took less time but made more errors on two of the three days, though the differences are rather small. This again would indicate that whatever the difference between punishment on the finger and the ankle may be, it is something which results in a slightly more cautious approach on the part of the ankle-punished individuals.

In Section 2 it was suggested that the average number of seconds per response was a fair index of the caution exerted by the subject. Using the formula presented in that section,

the average number of seconds per response was calculated for both the first trial and the later trials for each group each day and the results are presented in Table V*A*. In Table V*B* are given the coefficients of the reliability of the difference between the means.

TABLE V*A*

THE AVERAGE NUMBER OF SECONDS PER RESPONSE FOR (*a*) FIRST TRIAL, (*b*) LATER TRIALS, FOR EACH GROUP ON EACH DAY

	Day I	Day II	Day III
(*a*) First trial			
Ankle group	3.663 (±.177)	3.179 (±.157)	3.302 (±.193)
Finger group	3.574 (±.203)	2.943 (±.159)	2.812 (±.137)
Control group	2.652 (±.248)	2.665 (±.253)	2.362 (±.171)
(*b*) Later trials			
Ankle group	2.927 (±.221)	2.267 (±.129)	2.140 (±.111)
Finger group	2.753 (±.145)	2.098 (±.076)	2.121 (±.083)
Control group	2.610 (±.199)	1.850 (±.117)	2.096 (±.128)

TABLE V*B*

THE COEFFICIENTS OF THE RELIABILITY OF THE DIFFERENCES BETWEEN THE MEANS GIVEN IN TABLE V*A*

	Day I	Day II	Day III
(*a*) First trial			
Ankle-Finger	.33	1.05	2.07
Finger-Control	2.88	.93	2.05
Ankle-Control	3.32	1.73	3.65
(*b*) Later trials			
Ankle-Finger	.65	1.13	.12
Finger-Control	.58	1.78	.18
Ankle-Control	1.06	2.39	.26

In Table V*A*, part (*a*), it is clearly evident that the punished groups are slower than the control group in rate of response during the first trial. In Table V*B*, part (*a*), the differences between the means of the punished and the control groups are seen to be two and three times their probable errors on two of the three days. The smaller differences obtained on the second day may be the result of a chance factor, in that three members of the control group varied markedly from the mode of their group in such a way as to greatly increase the average of that group. In so far as the number of seconds per response is an index of caution, these results

further justify the tentative conclusion that punishment results in a more cautious approach to the problem during the first trial.

The consistent daily increase of the significance of the separation between the means of the finger and the ankle groups in part (*a*) of Table V*B* is interesting. It probably indicates that the rapid tapping of the finger-punishment individuals on the second and third day resulted in the avoidance of most of the unpleasant consequences of the electric shock and therefore permitted a faster rate of response. The equally significant separation between the finger and control groups probably indicates that, though the sensory consequences were not so unpleasant, the finger group still took time to memorize the combination during the first trial.

Section 6.—The effect of the punishment during the later trials:

The adaptation necessary after the completion of the first trial is, first, to retain the correct combination in the proper order, and, second, to correct any 'errors in order' (see footnote number 3) that may occur with as few trial and as little exploration as possible. As previously suggested, careful memorization in the first trial would decrease the number of later trials necessary to reach the criterion. In addition to this, a cautious attitude in the exploration following an unexpected 'error in order' would probably reduce the number of errors made in the later trials.

TABLE VI*A*

TIME AND ERROR AVERAGES IN THE TRIALS, AFTER THE FIRST, FOR EACH GROUP ON EACH DAY

	Day I	Day II	Day III
Time in seconds			
Ankle group	362.4	197.0	165.7
Finger group	377.2	223.9	163.3
Control group	399.8	174.0	210.5
Errors			
Ankle group	55.7 (± 6.68)	22.5 (±3.36)	18.5 (±3.03)
Finger group	79.8 (±13.97)	26.6 (±3.91)	17.1 (±2.87)
Control group	77.6 (± 9.56)	38.0 (±7.03)	38.0 (±5.04)

TABLE VI*B*

COEFFICIENTS OF THE RELIABILITY OF THE DIFFERENCES BETWEEN THE MEANS OF THE ERROR PART OF TABLE VI*A*

	Day I	Day II	Day III
Errors			
Ankle-Finger	1.56	−.79	.32
Control-Finger	−.13	1.41	3.61
Control-Ankle	1.88	1.98	3.31

Since the means of the time part of Table VI*A* show no consistent relation, no attempt is here made to evaluate the statistical significance of the separations between the means.

In the error part of Table VI*A* it is plainly evident that the punished groups made fewer errors in the later trials on the second and third day than did the control group. In Table VI*B* the reliability of the separation is seen to increase from day to day until the differences are more than three times their probable errors on the third day. This justifies the conclusion that punishment results in a reduction of errors, both by causing more careful memorizing in the first trial and by assuring a cautious attitude in the correction of 'errors in order' in the later trials.

In an earlier section the marked difference in errors between the finger-punished and ankle-punished groups on the first day was accounted for in terms of the differences in the behavior adaptations made by the two groups. A comparison of the first day averages in Tables IV*A* and VI*A* shows clearly that the difference in total errors is primarily due to the number of errors piled up by the finger-punished individuals in the *later* trials. Nine of the finger group made more than one hundred errors in the later trials, while only four of the ankle-punished group exceeded this number. This increase in errors in the later trials would suggest the hypothesis that the finger punishment exerted a disruptive effect only when it came unexpectedly in the course of the later trials while learning the first combination.

A comparison of the average time per response in parts (*a*) and (*b*) of Table V*A* shows that the rate of response

is much slower in the first trial than in the later trials with each group on each day. In Table VII are given the coefficients of the reliability of the differences between the means of the first trial and of the later trials.

TABLE VII

THE COEFFICIENTS OF THE RELIABILITY OF THE DIFFERENCES BETWEEN THE MEAN RATE OF RESPONSE OF THE FIRST TRIAL AND THE MEAN OF THE LATER TRIALS

	Day I	Day II	Day III
Ankle-Ankle	2.69	4.80	5.23
Finger-Finger	3.29	4.49	4.31
Control-Control	0.13	2.93	1.24

The fact that the coefficients of reliability of the punished groups are markedly greater than those of the control group indicates that the punishment has by far the greater effect in the initial trial. This interpretation is further justified by comparing the coefficients of the reliability of the differences in parts (*a*) and (*b*) of Table V*B*. Here again it is evident that there is much greater and more significant separation between the punished and the control groups in the first trial than in the later trials. This, then, justifies the conclusion that punishment as a factor producing cautious adjustment is far more effective during the exploration and memorization of the initial trial than during the later trials.

SUMMARY OF RESULTS

Section 1.—The variable effect of punishment upon the respective criteria of learning:

Errors averages were more definitely and consistently affected by the punishment than were the time or the trial averages.

Section 2.—The effect of punishment upon adaptation to a novel problem:

A. Punished vs. Control:

(1) The punished groups learn the combination on the first day with fewer trials and fewer errors, but require approximately the same total time to complete the learning.

(2) The number of seconds *per response*, or per key touched, as determined by the formula,

$$\frac{\text{Total Time}}{\text{Total Errors} + 10\,(\text{Total Trials})},$$

is used as an index of the degree of caution shown by the punished groups. The first day differences between the averages of the punished and the control groups with this criterion are about two times their probable errors.

B. Ankle vs. Finger:

(3) The ankle-punished group learned the problem on the first day with fewer errors than did the finger-punished group. This is interpreted as being more than a chance difference since, (*a*) there is a marked difference in the behavior adjustment of the two groups; (*b*) nine of the finger group made more than one hundred errors in the 'trials after the first', while only four of the ankle group exceeded this number.

Section 3.—The day to day adaptation to a problem involving punishment:

A. Punished vs. Control:

(1) There is no consistent relation between the punished groups and the control group in average time from day to day.

(2) The punished groups consistently learned with fewer trials and fewer errors. The relative size of the standard deviations indicates that there is less individual variability in the number of trials and errors with the punished groups.

B. Ankle vs. Finger:

(3) The finger-punished group consistently learned in less time than the ankle-punished group, but required more trials and made more errors on days I and II.

Section 4.—The relation of the intensity of the punishment to the criteria of learning:

Ankle vs. Finger:

(1) Practically no relation was found between the reported intensity of the punishment and the respective criteria within either of the punished groups. This is not to be considered as contrary to the findings of earlier experimenters,

since the problem, the scale of estimating, and the objective intensity of the punishment used have varied in each instance.

(2) There is, however, a day to day relation between the 'average number of seconds per response' and the average 'intensity reported' when the finger group is compared with the ankle group.

(3) The differences in average time, trials and errors are greatly reduced when the two punished groups are equated in average intensity reported. This indicates that the spatial factor is relatively unimportant after the first day when intensity is kept constant.

Section 5.—The effect of punishment during the exploratory trial:

A. Punished vs. Control:

(1) The punished groups required more time but made fewer errors than did the control group in the first trial on each day.

(2) Using the 'number of seconds per response' as a criterion of caution, the punished groups were found to be much more cautious in the first trial than was the control group.

B. Ankle vs. Finger:

(3) The finger group took less time on all three days, but made more errors than did the ankle group in the first trial on two of the three days.

(4) By the criterion used in (2) above, the ankle group is more cautious than the finger group on all three days.

Section 6.—The effect of punishment during the *later* trials:

A. Punished vs. Control:

(1) With the exception of the finger group on the first day, the punished groups consistently made fewer errors in the later trials than did the control group. The differences became more marked on the successive days. This indicates that the punished groups were more efficient in learning in the first trial on the successive days.

(2) Again using the number of seconds per response, the punished groups which were consistently and significantly

more cautious in trial one were found to be only slightly more cautious than the control group in the later trials.

B. Ankle vs. Finger:

(3) There are no consistent differences between the ankle and the finger groups in the time and error records during the later trials.

Conclusions

I. Punishment Compared with No Punishment

(1) The electric shock of the intensity here used resulted in a more cautious attitude on the part of the learners in this experiment.

(2) Caution on the part of the punished learners resulted in their learning the combinations used with fewer trials and fewer errors than the non-punished learners, though they required the same amount of time.

(3) An analysis of the data of the first trial as compared with the later trials shows that the punished groups were significantly more cautious in the exploratory initial trial than in the later trials.

II. Ankle Punishment Compared with Finger Punishment

When the intensity of the punishment is kept constant, the factor of spatial separation seems to be of minor importance in the learning of a problem of this type.

References

1. Brown, W., and Whittell, F., Yerkes multiple choice method with human adults. *J. Comp. Psychol.*, 1923, **3**, 305–318.
2. Bunch, M. E., The effect of electric shock as punishment for errors in human maze learning. *J. Comp. Psychol.*, 1928, **8**, 343–359.
3. Johanson, A. M., The influence of incentive and punishment upon reaction time. *Arch. Psychol.*, 1922, **8**, No. 54, pp. 53.
4. Rexroad, C. N., A continuous multiple choice reaction apparatus. *J. Exper. Psychol.*, 1925, **8**, 325–336.
5. Rexroad, C. N., Administering electric shock for inaccuracy in continuous multiple-choice reactions. *J. Exper. Psychol.*, 1926, **9**, 1–18.
6. Vaughn, J., *Positive versus negative instruction.* Publications of the National Bureau of Casualty and Surety Underwriters, Educational Series, Volume II, pp. 172.
7. Vaughn, J. and Diserens, C. M., The relative effects of various intensities of punishment on learning and efficiency. *J. Comp. Psychol.*, 1930, **10**, 55–66.

8. WATSON, J. B., Recent experiments on how we lose and change our emotional equipment. *Ped. Sem.*, 1925, **32**, 349–371.
9. WONG, HISHING, AND BROWN, W., Effects of surroundings upon mental work as measured by Yerkes' multiple choice method. *J. Comp. Psychol.*, 1923, **3**, 319–330.
10. YERKES, R. M., A new method of studying the ideational behavior of mentally defective and deranged as compared with normal individuals. *J. Comp. Psychol.*, 1921, **1**, 369–394.

(Manuscript received September 2, 1930)

LETTER-POSITION EFFECTS, IN THE RANGE OF ATTENTION EXPERIMENT, AS AFFECTED BY THE NUMBER OF LETTERS IN EACH EXPOSURE

BY H. R. CROSLAND

University of Oregon

In 1928 Miss Georgia Johnson and I published [1] a paper which, though directed at a wholly different problem, brought to light the fact that the letter-positions, from left to right on the various cards of exposure, became in recall less and less effective towards the right. We had presented through the Dodge mirror-tachistoscope eight series of cards, for an exposure period of 150 σ, to 32 subjects. The series were composed of cards carrying from 3 to 10 letters per card, each series containing five examples of its number of letters. The results when recorded, tabulated, and analyzed revealed a curve of descending effectiveness of recall for the 10 positions from left to right. This curve, however, was constructed by pooling together the results from all series, the short as well as the very long series. The curve and its basal numerical data have, so far, stimulated [2] two major criticisms: (1) The composite nature of the curve obscures the effects of position in each individual series, so that the results from the very short series, recalled with almost complete efficiency, would contaminate the results from the very long series by giving too high values for the first four or five letter-positions. At any rate, it is desirable to envisage each series, short as well as long, separately. And (2) we erred in not using series uniform in length and each having a number of letters above the average subject's range, it being argued that the very short series are recalled very uniformly, with almost 100 percent efficiency, so that the three letter-positions in it would heavily

[1] The range of apprehension as affected by inter-letter hair-spacing and by the characteristics of individual letters, *J. Applied Psychol.*, 1928, **12**, 82–124.

[2] These two criticisms have been projected by Dr. M. A. Tinker, *Psychol. Bull.*, 1929, **26**, 227f. and 611–613.

weight any series containing more than an average range number of letters. In other words, the curve of 1928 should have begun its descent at a lower level and should have fallen throughout its course more slowly than it did. In addition to the above criticisms, there are three other possible objections: (*a*) the very objective nature of the eight series was such as to present a descending frequency of letters in positions one through ten; thus, the curve merely paralleled the curve of the frequency of the objectively arranged letters. Elsewhere [3] this assumed criticism has been answered in detail and may be assumed to be worthy of no further consideration. (*b*) Uncontrolled fixation of the subjects' lines of regard may have resulted in a left (habitual-reading) type of fixation, thus resulting in the left-to-right reductions in position efficiency. Had fixation been effected and maintained at the right, in each exposure, our curves would, it is argued, have been reversed in their declinations. And (*c*) the left-to-right reduction in efficiency may have resulted from the priority value of memory, since most subjects recall from left to right and write down their reproductions in a similar manner.

To meet criticisms (1), (2), and (*b*), above, I have repeated the experiment.[4] Criticism (*c*) was not before me at the moment when the duplication experiment was planned so that it must await a testing out in the future. The experiment as planned did not envisage objection (*b*), above, but the newer technique and the results as obtained are nevertheless applicable.

A brief description will present the details of the duplication experiment; comparisons of the earlier with the later techniques will be made in parallel columns; and then will be presented the obtained results with some discussion of their significance. 140 [5] sophomore, Beginning Psychology Laboratory, students served as subjects. The experiment, under close supervision, occurred as a regular laboratory exercise.

[3] *Psychol. Bull.*, 1929, **26**, 375–377.

[4] Failure to repossess the original data of Miss Johnson's experiment prevented the reworking of them; hence, the entire experiment was duplicated, with a few minor changes in technique.

[5] Several records, incomplete or otherwise inaccurate, have been deleted.

Two subjects worked alone in a single room, one serving as experimenter and the other as observer. After the first trial (of two) with the 60 cards of the experiment, they exchanged functions, and continued to alternate until the two trials had been had by each of them. Each pair of subjects worked in a small room illuminated by a dome light of 75 Watts and by a window at the observing subject's right. The rooms, situated on the north side of the building, gave indirect sunlight, possibly a trifle less in the afternoon than in the morning; and the class was equally composed of morning and afternoon sections, approximately equal in enrollment. The presentations were made by a gravity-propelled, falling-shutter, type of tachistoscope, the card bearing the letters being held stationary with a presentation, open-window passing in front of the card. The tachistoscope was 77 cm in height, and its card was situated 38 cm above the table top on which it stood vertically. The exposure slit in the falling screen was 5.5cm by 8.5 cm in area. This slit and the exposure card became coincident, with the falling of the shutter, for a period of 100 σ. The subject faced the tachistoscope at a distance of 30 to 40 cm from the card, presented at the level of his eyes. He sat in a comfortable, straight chair, just at the edge of the table on which rested the tachistoscope. 'Ready' signals always directed his gaze to the center of the exposure area just prior to a presentation. The exposure material or printed card was affixed in a stationary position in a special container 1 cm behind the falling shutter and approximately 1.5 cm behind the opening in the screen. The disengaging of a hinge at the top of the apparatus permitted the card-board shutter to fall, thus bringing its opening coincident both with the front opening and the exposure card behind for exactly 100 σ. The exposure materials consisted of 60 white cards, 88 mm square, carrying in black print the 26 letters of the alphabet arranged in random groupings. The 60 were divided into six series of 10 cards each, each series being characterized by its number of letters, such as the 4's, 5's, and so on through the 9's. The print employed was Caslon Bold capitals in size 18-point. Each card possessed both on front and back a number-symbol

by which its records could be controlled and properly segregated. In the two trials by each subject different random orders of the cards were used and no two subjects employed the same order of presentation. The alternation, as between experimenter and subject, served to prevent familiarity from proving to be a very deleterious factor in any subject's records.

The recording was effected by the subject himself. He was supplied with four recording and tabulation sheets, already mimeographed for the purpose. On two of these (one for each trial) he merely noted down, in fixed positions, the letters he recalled after each presentation. Later, after completion of the two trials, he also wrote on these forms the letters as objectively correct. This gave him a basis for the studying of error, for tabulation, and for computation. The remaining two blanks afforded a means for scoring errors, both by letter position and by letter-series, and for differentiating the scores on the basis of letters correct in every respect and letters merely misplaced. The latter scores will not engage us in this paper.

The technique, with its likenesses and differences with that of 1928, may be comprehended by the parallel column comparisons as given on page 481.

The remainder of the paper will present results and conclusions based thereon. The results themselves will appear as tables and graphs from which several interesting deductions and cross comparisons will be drawn.

Results

Reliability

One object in running two trials in this experiment was to furnish a basis for computing the reliability of the investigation. This method of effecting self-correlation, for reliability, became necessary when it was decided that no two subjects would receive the same order of presentation of the 60 cards and when the decision was formulated to analyze the results both by letter-position and by series-length. Any such plan, it is obvious, makes the usual odd-even, half-half, correlation

Items of technique	Crosland and Johnson's experiment of 1928	The 1930 duplication of the 1928 experiment
(1) Exposure-apparatus..	Dodge, mirror-tachistoscope	Home-made, drop, gravity tachistoscope
(2) Exposure-time.....	150σ	100σ
(3) Distance from eyes to exposure-field...	12 cm from diagonal transparent mirror; 46 cm from exposure-cards	30 to 40 cm
(4) Height of exposure...	On eye-level	On eye-level
(5) Lighting of room....	Every subject worked singly in same room, with specially controlled lighting	Two subjects at a period, in a single room; many rooms used; lighting uncontrolled; one north window to each room
(6) Fixation-point.....	Center of field	Center of field
(7) Size of exposure-card.	88 sq mm	88 sq mm
(8) Size of type.....	10-point, lower case	18-point, all capitals
(9) Font of type.....	Body Caslon	Caslon Bold
(10) Number of series....	8	6
(11) Number cards per series.....	5	10
(12) Number letters per card.....	3 to 10	4 to 9
(13) Total number of cards	40	60
(14) Total number of letters by positions per subject.....	40 40 40 35 30 25 20 15 10 5 1 2 3 4 5 6 7 8 9 10	60 60 60 60 50 40 30 20 10 1 2 3 4 5 6 7 8 9
(15) Practice.....	Before each exposure, 4 trial exposures	None
(16) Trials (presentations of set of 40 cards)..	1	(Presentations of set of 60 cards) 2 (trials 1 and 2)
(17) Order of cards.....	Chance	Chance (reshuffled for each trial)
(18) When recorded.....	Immediately after each exposure	Immediately after each exposure
(19) Manner of recording.	By subject, by exposure no., on blank paper	By subject, by presentation no., on prepared blanks
(20) By whom corrected and scored.....	Experimenter	Subject, under supervision (checked by experimenter).
(21) By whom scored and tabulated for letter-positions.....	Experimenter	Recorded on prepared chart, by subject, after experiment was completed, rechecked by experimenter
(22) Control of the order of written reproduction.....	None, subject being free to reproduce in any order but required to note all blanks in recall by blanks on record pad	None, subject being free to reproduce in any order but being required to denote blanks in recall by blanks on record pad
(23) Chin-rest and eye-hood.....	Yes	None
(24) Number of subjects..	32	131 (plus 9 deleted for inaccuracies)
(25) Consideration given effects of letter-position.....	Such results considered as by-products of a larger purpose	The prime and sole purpose, omitting reference to transpositions, spacing, etc.

impossible. Hence, we have correlated the scores for letters entirely correct (correct and properly placed) in the first trial against the second trial. This has been done, first, for each letter-position (1 through 9, from left to right) irrespective of length of series (all series being pooled); and, second, for each series irrespective of letter-position (all positions being pooled). Table I presents the two sets of reliability coefficients.

TABLE I

INTER-TRIAL CORRELATION (RELIABILITY) COEFFICIENTS

A. By Letter-Position

Positions (1–9)

1	2	3	4	5	6	7	8	9
.77±.02	.80±.02	.80±.02	.79±.02	.76±.02	.81±.02	.87±.016	.89±.013	.79±.02

Mean = .81

B. By Series (Lengths)

Series (4's–9's)

4's	5's	6's	7's	8's	9's
.65±.03	.76±.026	.67±.03	.72±.029	.71±.03	.70±.03

Mean = .70

Analysis of the foregoing table (I) reveals the fact that reliability, regarded from any point of view, is ample to meet Kelley's minimum requirement of .4 for group results from a population in excess of 30 subjects. The lowest coefficient, .65, is obtained in the case of the shortest series, the 4-letter cards; and the highest, .89, is obtained for the 8th letter-position from left to right. Contrary to expectations, the higher mean reliability, .81, characterized the letter-position scores, as against a mean reliability of .70 for the six series of different lengths of cards. Dividing the letter-positions into left- and right-hand halves and computing the mean r for each half, we obtain .784 for the left-hand positions and .824 for the right-hand positions. This is a result reversed from the one expected because of the allegation that greater variations between subjects and between the different reactions

of any one subject at different times should accumulate with greater frequency in the longer series, the series determining almost exclusively the scores for the last four letter-positions. Obtaining, likewise, mean reliabilities for the two halves of the series-lengths, we discover .69 for the shorter series and .71 for the longer, thus affording a difference which, though small, accords with that above: reliability is greater for the longer series and for the right-hand letter-positions. For the experiment in its entirety reliability is sufficiently high to attach significance to the results which are to follow. Lest this conclusion be challenged by asserting that this high reliability has, in part, resulted from the learning characterizing the second trials, as against the first, it will be wise to examine the effects of learning more closely, although it is reasonable to suppose that learning would not affect all 131 subjects alike, and would therefore reduce reliability rather than raise it. This examination may be done by scrutiny of Table II.

That there was considerable learning is unmistakable. Regarding the means of all letter-positions, from Table II*A* the gain from the first trial was approximately .51 or 11.3 percent, while for the left-hand letter-positions alone it was 4.25 or 19.58 percent and for the right-hand positions alone it was .46 or 5.67 percent. Similarly, consideration of the means of all series (Table II*B*) reveals in the second trial a gain of .56 or approximately 15 percent, and the three shortest series exhibit a gain of .82 or 14.26 percent while the three longest series contained a gain of .48 or 17.0 percent. We may conclude, therefore, that there was no appreciable amount of learning in the right-hand letter-positions, and hence, for those scores reliability was materially increased; and that the appreciable learning in the left-hand positions was a factor in the lower reliability there shown. We may conclude further that learning was almost equally distributed over long and short series, being a trifle greater in the longer series, and being considerable it served to reduce reliability to a point well under that for the letter-positions. Being thus reasonably assured that the coefficients of reliability,

TABLE II

A. THE EFFECTS OF LEARNING ON LETTER-POSITION MEAN SCORES WHERE ALL SERIES HAVE BEEN POOLED TOGETHER

Positions	1	σ	2	σ	3	σ	4	σ	5	σ	6	σ	7	σ	8	σ	9	σ	Mean–Mean of 39 items
Trial 1	51.89	8.06	35.23	11.38	26.7	10.28	18.6	9.03	7.49	6.02	3.31	4.43	2.22	3.28	1.44	2.88	0.97	1.79	3.79
Trial 2	54.1	7.11	40.9	10.39	32.4	11.45	24.4	10.42	9.34	6.57	3.99	4.98	2.15	3.54	1.39	2.96	0.86	1.81	4.35
Diff.	+2.21		+5.7		+5.7		+5.8		+1.85		+0.68		−0.07		−0.05		−0.11		+0.56
% of gain	4.3		16.2		21.5		31.2		24.7		20.5		−3.15		−3.4		−10.3		15.00
Diff./σd	4.71		9.24		9.24		10.09		4.78		2.57		0.44		0.13		0.33		
Chances	100		100		100		100		100		99.4		67		55		63		

B. THE EFFECTS OF LEARNING ON SERIES MEAN SCORES WHERE ALL LETTER-POSITIONS HAVE BEEN POOLED TOGETHER

Series	4's	σ	5's	σ	6's	σ	7's	σ	8's	σ	9's	σ	Mean–Mean of 39 items
Trial 1	30.16	6.60	28.45	8.30	25.02	8.13	22.68	6.70	22.08	8.07	19.53	7.08	3.79
Trial 2	34.16	6.10	32.55	8.02	28.80	8.80	26.25	8.32	25.68	8.00	23.22	8.25	4.35
Diff.	+4.00		+4.10		+3.78		+3.57		+3.60		+3.69		
Diff./σd	8.43		8.11		6.00		6.88		6.42		6.79		
Chances	100		100		100		100		100		100		
Mean diff.	+1.00		+0.82		+0.63		+0.51		+0.45		+0.41		+0.56
% of gain (by mean diff.)	13.26		14.41		15.11		15.74		16.30		19.00		15.00

lowered rather than raised by such learning as occurred between trials one and two, are satisfactorily high for the ascertainment of group differences, we may now turn directly to the detailed group results themselves.

The total scores and the mean scores, for each of the two trials and for both trials, arranged by series and the letter-positions may be seen in Table III. Graphs constructed

TABLE III

TOTAL AND MEAN SCORES FOR 125 AND 128 SUBJECTS, BY LETTER-POSITIONS, BY SERIES, AND BY TRIALS

First Trial Totals (N = 128)

Series	Letter-Position: 1	2	3	4	5	6	7	8	9	Mean
4's	1183	986	832	858						965
5's	1171	883	640	461	486					728
6's	1152	768	614	282	154	236				534
7's	1053	697	549	253	109	80	164			415
8's	1079	629	463	286	108	67	76	121		354
9's	1015	540	323	251	104	51	43	65	125	280
Mean	1109	750.5	570	398.5	192	108.5	94.3	83	125	

Second Trial Totals (N = 125)

										Mean
4's	1200	1067	912	963						1035
5's	1188	975	763	579	563					814
6's	1150	888	750	391	160	265				601
7's	1081	812	628	369	156	91	150			484
8's	1101	754	584	391	137	59	70	117		402
9's	1056	630	415	365	152	86	49	57	108	324
Mean	1129	854	675	509.7	233.6	125	89.7	87	108	

Mean Totals of Both Trials (N = 125)

										Mean
4's	1177	1017	863	900						989
5's	1163	925	693	513	525					764
6's	1138	813	675	338	155	248				561
7's	1055	746	582	308	131	85	155			438
8's	1077	684	518	335	121	62	72	118		375
9's	1023	579	365	305	127	68	46	60	115	299
Mean	1105.5	794	616	449.8	211.8	116	91	89	115	

First Trial Means ($N = 128$)

										Mean
4's	9.24	7.7	6.5	6.7						7.54
5's	9.15	6.9	5.0	3.6	3.8					5.69
6's	9.0	6.0	4.8	2.2	1.2	1.8				4.17
7's	8.2	5.4	4.3	2.0	0.85	0.6	1.3			3.24
8's	8.4	5.0	3.6	2.2	0.84	0.52	0.59	0.94		2.76
9's	7.9	4.2	2.5	1.9	0.8	0.39	0.33	0.5	0.97	2.17
Mean	8.65	5.87	4.45	3.1	1.5	.83	.74	.72	.97	3.79

Second Trial Means ($N = 125$)

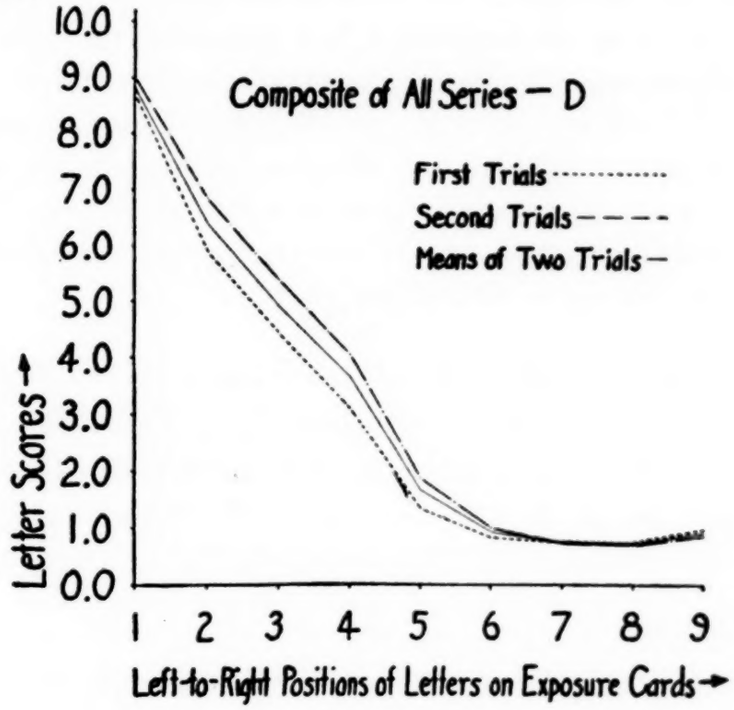
										Mean
4's	9.6	8.54	7.3	7.7						8.54
5's	9.5	7.8	6.1	4.63	4.5					6.51
6's	9.2	7.1	6.0	3.1	1.3	2.1				4.80
7's	8.6	6.5	5.0	3.0	1.25	0.73	1.2			3.75
8's	8.8	6.0	4.7	3.1	1.09	0.47	0.56	0.93		3.21
9's	8.4	5.0	3.3	2.9	1.2	0.69	0.39	0.46	0.86	2.58
Mean	9.02	6.82	5.40	4.07	1.87	1.00	.72	.70	.86	4.35

Mean Totals of Both Trial Means ($N = 125$)

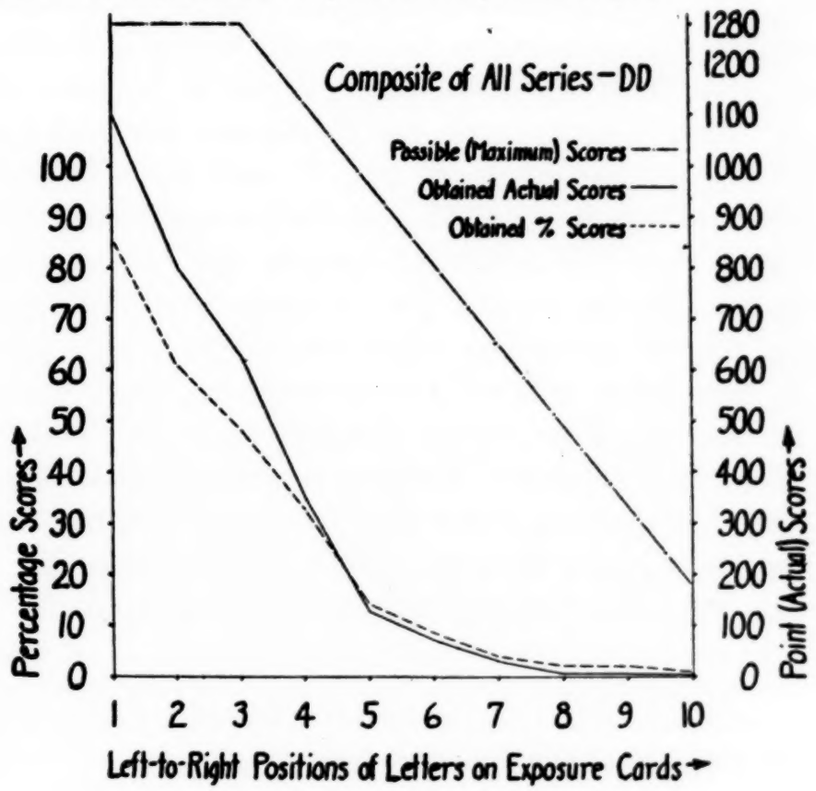
										Mean
4's	9.42	8.14	6.9	+7.2+			+ + = wrongly			7.92
5's	9.3	7.4	5.54	4.1	4.2		graphed as 7.6 in			6.11
6's	9.1	6.5	5.4	2.7	1.2	2.0	Graph *C*			4.48
7's	8.4	6.0	4.7	2.5	1.05	0.68	1.2			3.22
8's	8.6	5.5	4.1	2.7	0.96	0.49	0.57	0.94		2.98
9's	8.1	4.6	2.9	2.4	1.0	0.54	0.36	0.48	0.92	2.37
Mean	8.82	6.36	4.92	3.60	1.68	.93	.71	.71	.92	4.07

from their pooled data, without distinction between series, are shown in Fig. *D*. Figure *DD* reproduces the composite graphs of 1928. Note the striking similarities between them. Drawn from the same scale, in the vertical axis, they would superpose themselves with great approximation. The main differences between them consist in a slightly greater declination of the 1928 graph and in its closer approach to zero scores for letter-positions 8, 9, and 10, a result which can be charged to its employment of 10-point type. Another difference is the flatter zone for positions 2 to 4 in the present graphs (*D*) than in the earlier ones. It is a strange coincidence that in both there are breaks, and in the same direction of declination, at positions 2, 4, and 5. In other words, the same composite results have been yielded in two different investigations,

Number of letters correctly recalled in each letter-position. Mean total scores for 125 subjects by trials.



Number of letters correctly recalled in each letter-position. Total scores for 32 subjects in first trial (Crosland and Johnson's Experiment, 1928).



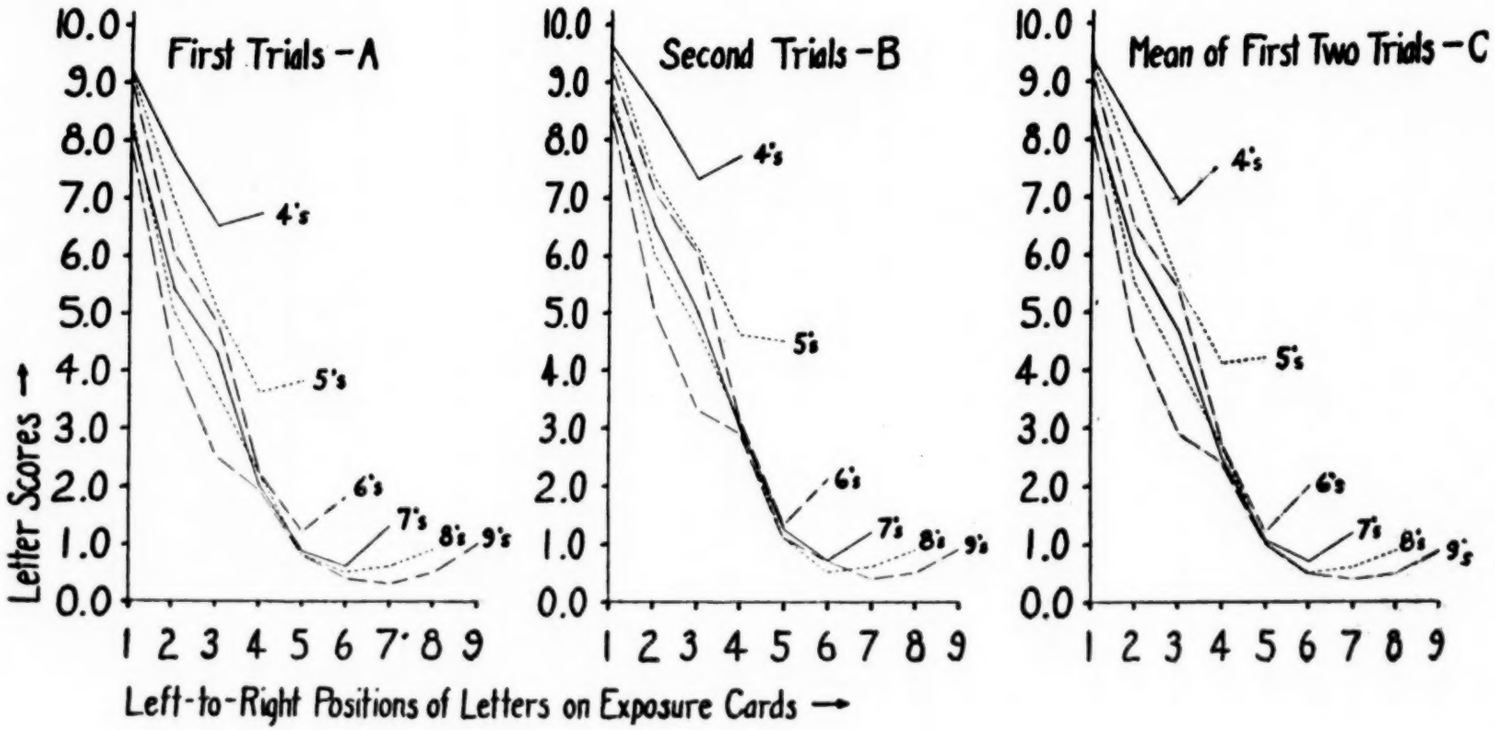
separated by two years in time and effected on different subjects: on one group of 32 and on another of 125. This coincidence and confirmation has occurred in spite of any alleged clumsiness of the second investigation and in the face of any and all inaccuracies which may have characterized the scorings and tabulations effected by sophomore subjects. In short, we have here apparently a real result, an incontrovertible fact. It remains, however, to apply criticisms (1) and (2) to this set of composite graphs.

Individual Series Results (All Cards Considered)

To apply criticism (1) to these data, Table III, we shall graph them by separate series and separately by trials, as shown in Figs. *A*, *B*, and *C*. One will note at a glance that there are striking similarities between the curves for the 4's, 5's, 6's, 7's, 8's, and 9's in each figure; that the greatest similarity exists between the 6's, 7's, 8's, and 9's; that the 5's are not very dissimilar to them; and that the 4's deviate most greatly from all others. One will note also that the 4's drop by approximately 40° below the horizontal, the 5's by 50° and the others by approximately 55°. 15 degrees' difference is the maximum, as between the 4's and the 9's, in declination, and there are only 10 degrees of difference between the 5's and 4's. It is likewise obvious that great similarities exist between the graphs of *A*, *B*, and *C*, and both *D* and *DD*. The differences are not alarming in their proportions, although it is a fact that the composite graphs did obscure the differences between the 4's and 5's. Graphs *D* and *DD* revealed an equalizing or averaging effect for the first four positions (1 to 4), instead of inflated values for them, since the longer series reduced the high values afforded these positions by the shortest series (the 4's).[6] Neither statement, that the short series raised the values of the first four positions, nor that the longer series reduced their values, is wholly justified. What happened was that Graphs *D* and *DD* presented, for these

[6] It is regrettable that the second investigation could make no use of a 3-letter series. The cards, formerly prepared for the Beginners' Laboratory, were used just as they existed. Time did not permit the addition of an extra series to make the later investigation wholly comparable with that of 1928.

Number of letters correctly recalled in each series—10 cards for each letter-position—The mean total scores for 125 subjects by trials.



positions, merely the mean results of the short and long series. The decomposing of Graph *D*, into *A*, *B*, and *C*, has brought forth one result, wholly hidden in the earlier investigation, which is deserving of attention. Every curve, in *A*, *B*, and *C*, turns upward at its right-hand end. In other words, the final character is more effective for recall than the next to last character. The longer the series becomes the more does this fact apply itself to the second from last, and to the third from last, character, etc. This slight increase in efficiency at the farthest right side is not, however, great enough to cancel the reduction (in efficiency in the final character) brought about by the increase in length of series. To demonstrate this, one may compare the elevated right-end of the 9's with that of the 8's, and so on back to the 4's. Nor is this augmented attention value, of final right character, great enough to offset the precipitous and then more and more gradual left-to-right declination in efficiency exhibited in each curve in all figures presented above.

Percentages of Cards Recalled Without a Single Error (by Letter-Positions and by Series)

Criticism (2) we may prepare to answer by discovering how many cards, in each series, were recalled without error, for it is these which are supposed to do such damage to the scores for the letter-positions by raising the values obtained from the longer series. Table IV contains these percentages and Fig.

TABLE IV

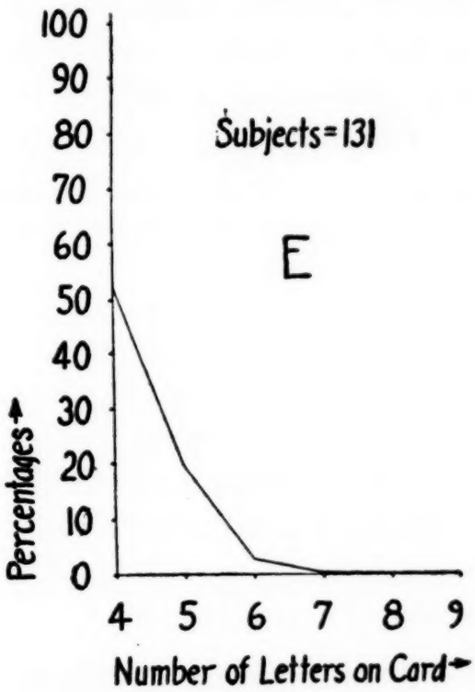
PERCENTAGES OF CARDS, BY SERIES, WHICH WERE RECALLED WITHOUT ERROR

Series	4's	5's	6's	7's	8's	9's
Percent	52.1	19.9	2.7	0.3	0.0	0.0
Containing one or more errors	(47.9)	(80.1)	(97.3)	(99.7)	(100.0)	(100.0)

E presents them graphically. In fully one-half of the cards in the shortest series one or more errors occurred. Certainly in Series 6, 7, 8, and 9 the cards recalled without error were infinitesimally small in frequency and could do no possible

damage by contaminating, with perfect position scores, the composite or the individual series graphs. 80 percent of the 5's, a comparatively short series, were recalled with one or more errors per card. Certainly 80 percent is a large enough number of cards to weight heavily positions 1 to 5 with scores

Percentage of all cards (1310) recalled without errors in each series.



taken from cards containing error and thus lying at or above range threshold. 48 percent of cards in Series 4 is a respectable number of cards, representing exactly 627 of 1310 cards. No one can justly say, therefore, that the positions 1, 2, 3, and 4 are weighted very heavily by perfect scores for these positions from Series 4. In fact, the suspicion seems justified that it was wise to employ both long and short series; that an arbitrary determination of range, at say 5, with the consequent elimination of all cards carrying less than six letters, would have been a most grievous mistake and at best a most arbitrary and *a priori* procedure.

It is perhaps a mere coincidence that Graph *E* is shaped somewhat as Graphs *D*, *DD*, *A*, *B*, and *C*, although its early declination is sharp and tapers off to almost ɔo further descent at Series 7.

Composite Scores and Graphs, Taken from Cards Containing One or More Errors, by Letter-Positions

Following our procedure of 1928 and of the first part of this paper, we may approach more closely Criticism (2) by graphing in composite fashion the error scores effected from cards only containing one or more errors, thus throwing out all contaminating, perfect, recalls, alleged to be so characteristic of short exposure-series. Table V presents such composite scores and Fig. *d* presents their graphic picture.

TABLE V

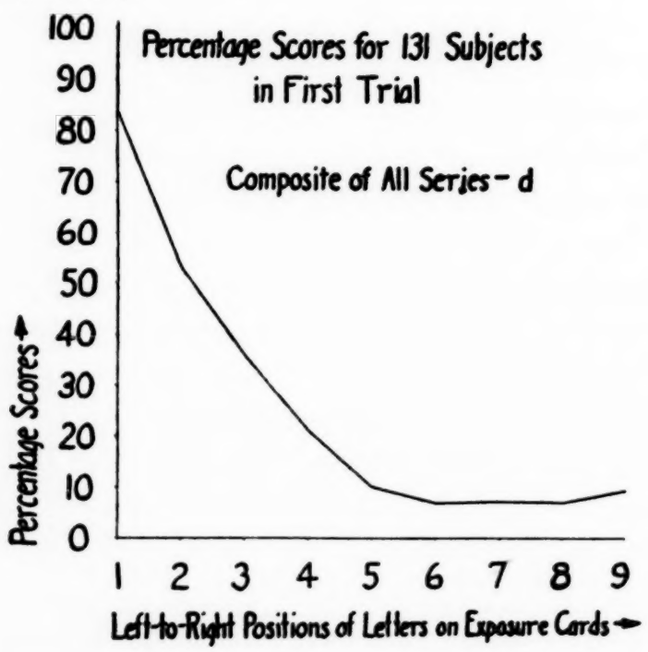
PERCENTAGE SCORES, FROM CARDS CONTAINING ONE OR MORE ERRORS, BY LETTER-POSITIONS (SERIES POOLED TOGETHER)

Letter-Positions

1	2	3	4	5	6	7	8	9
84.0	53.0	36.0	21.0	10.0	7.1	7.3	7.2	9.6

If one cares to compare Graph *d* with *D* and *DD*, one will discover that the greatest effect produced by the elimination

Number of letters correctly recalled in each letter-position for only cards containing one or more errors.



of a few correct cards has been to change that region of the curve, between 2 and 5, from a convex to a concave one.

Positions 1 and 2 have been altered very little, if any, relatively speaking. Positions 5 to 9 are relatively unaffected. The whole curve is slightly depressed, but in an almost unobservable degree. In fact, it is merely positions 3 and 4 that are affected, and this modification has produced no more of an effect than to straighten out the bulge in the previous composite graphs. The essential results are evident in either graph, *D* or *d*; namely, there is a rapid loss in effectiveness of position as one moves from left to right across exposure cards. No decomposing of composite graphs and no eliminations of correct recalls have been sufficient to remove or even to alter that basic and outstanding fact.

Individual Series Scores and Graphs, from Cards Containing One or More Errors, by Letter-Positions

In support of the preceding paragraph we may now examine the data from which have been excluded all cards recalled with 100 percent efficiency, series by series. Such data comprise Table VI, and their graphs appear as Fig. *a*.

TABLE VI

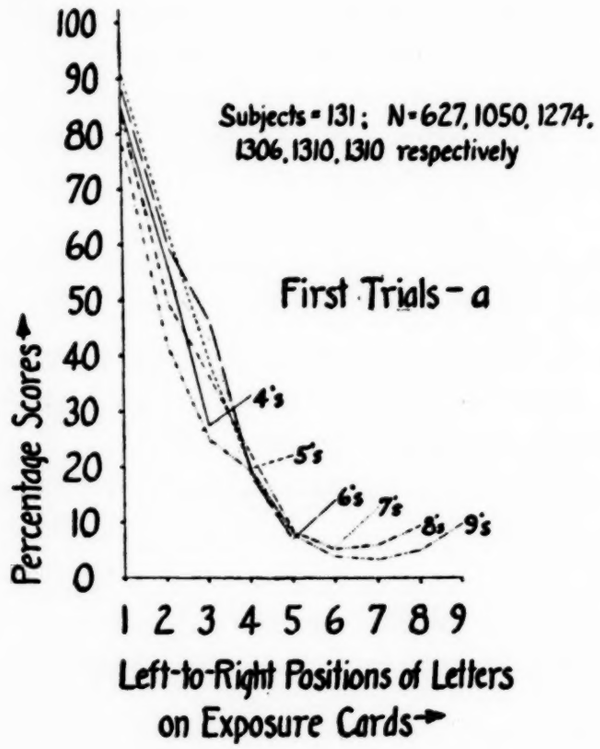
PERCENTAGES OF LETTERS (ON CARDS CONTAINING ONE OR MORE ERRORS) RECALLED CORRECTLY FOR EACH LETTER-POSITION IN EACH SERIES

Series	Positions								
	1	2	3	4	5	6	7	8	9
4's.....	84.4	56.5	27.4	32.7					
5's.....	90.9	62.6	38.7	19.6	22.2				
6's.....	88.6	60.2	45.8	18.8	7.0	14.1			
7's.....	81.9	53.8	42.8	19.8	8.2	5.7	12.7		
8's.....	84.0	50.0	36.0	22.0	8.4	5.2	5.9	9.4	
9's.....	79.0	42.0	25.0	19.1	8.0	3.9	3.3	5.0	9.7

Examination of Graph *a* reveals that all curves, 4's, 5's, and so on through the 9's, approach with remarkable exactitude almost complete superposition. This nearly complete superposing of the curves for the various series means not only that the angles of declination are the same throughout their courses but that the scores, for any one letter-position, as between series, are almost identical. To eliminate any

supposition that the rejection of all 100 percent cards has alone been responsible for this remarkable parallelness and near-superposition, we may compare the first-trial scores before and after the above-mentioned selection of cards containing one or more errors, in Table VII.

Percentage of letters recalled for each letter-position in each series on cards containing one or more errors.



The foregoing table (VII) is conclusive in demonstrating no change, with the deletion of all perfectly-recalled cards, for letter-positions 1, 6, 7, 8, and 9. Position 1 is one of the first five positions which were supposed to have given high scores because of the perfect scores of the short series, but when such perfect scores are eliminated, position 1 continues to maintain its highest rank as the most efficient reading position, left to right. Nor is the reduction effected by the elimination in scores for position 2 and 5 very appreciable. Hence, we may conclusively assert that of the first five positions, positions 1, 2, and 5 are unimpaired in rank-positions of effectiveness by reason of the deletion of the contaminating scores of the perfect reproductions of the short

TABLE VII

(Comparison of Item 4, Table III, with Table VI)

Comparative Likenesses and Differences Between Scores Before and After the Elimination of All Perfectly Recalled Cards

Series	Letter-Positions										
	1	2	3	4	5	6	7	8	9		
4's	92.4 84.4	77.0 56.5	65.0 27.4	67.0 32.7							Before After
	−8.0	−20.5	−37.6	−34.3	Mean = −25.1						Diff.
5's	91.5 90.9	69.0 62.6	50.0 38.7	36.0 19.6	38.0 22.2						Before After
	− .6	− 6.4	−11.3	−16.4	−15.8	Mean = −10.1					Diff.
6's	90.0 88.6	60.0 60.2	48.0 45.8	22.0 18.8	12.0 7.0	18.0 14.1					Before After
	−1.4	+ .2	− 2.2	− 3.2	− 5.0	−3.9	Mean = −2.6				Diff.
7's	82.0 81.9	54.0 53.8	43.0 42.8	20.0 19.8	8.5 8.2	6.0 5.7	13.0 12.7				Before After
	− .1	− .2	− .2	− .2	− .3	− .3	− .3	Mean = −0.23			Diff.
8's	84.0 84.0	50.0 50.0	36.0 36.0	22.0 22.0	8.4 8.4	5.2 5.2	5.9 5.9	9.4 9.4			Before After
	0	0	0	0	0	0	0	0	Mean = 0.0		Diff.
9's	79.0 79.0	42.0 42.0	25.0 25.0	19.0 19.0	8.0 8.0	3.9 3.9	3.3 3.3	5.0 5.0	9.7 9.7		Before After
	0	0	0	0	0	0	0	0	0	Mean = 0.0	Diff.
Mean diff.	−1.7	− 4.5	−8.55	− 9.0	− 4.2	−1.05	−0.1	−0.0	−0.0		

True Mean of Series Means = −4.3

series. There remains, then, the question of why positions 3 and 4 should show larger reductions by reason of the elimination. Scrutiny of the Graphs *A* and *a* reveals the turned-up or ascending, final ends of the curves of relative effectiveness of letter-positions in the individual series, namely, the fact that the final character in each series possessed an attention value in excess of that of its adjacent left-hand neighbor. In Graph *a* this fact is even more in evidence. Hence, when

letter-positions occupying these terminal points in the shorter series are thrown together with the scores for like letter-positions in the longer series (where naturally the curves are still in their descending slopes) the shorter series, by reason of the ascents in the terminal positions, do weight too heavily such position scores; but when all 100 percent recalls are rejected, since they are most numerous in the shorter series, and since in each of the shorter series the remaining cards show ascending error toward the right, the elimination disproportionately cuts down the scores for letter-positions 3 and 4. In other words, three factors are at work determining the values of the scores for positions 3, 4, and 5 in the composite graphs: (1) the general tendency (and the strongest) of decreasing effectiveness from left to right; (2) right-terminal attention value, having in composite graphs its largest effect upon the combined scores; and (3) the larger percentage of 100 percent recalls, in Series 4, hence equal scores between positions 1, 2, 3, and 4, which when rejected have, in conjunction with the two preceding factors, its largest effects in positions 3 and 4. Moreover, these factors functioned in importance in the order in which we have just presented them.

Further consideration of Table VII reveals the largest reduction, by series, in Series 4 and 5. The reduction is negligible in all other series, and is more than twice as large in Series 4 as in Series 5. For all series, considered together by their mean, the reduction in scores effected by the elimination of perfect cards, was almost negligibly small, amounting to only 4.3 percent.

Finally, attention should be redirected to Graph *a*. The largest change has consisted of the forcing of the 4's and 5's into parallelness with the 6's, 7's, 8's, and 9's. All curves have been more or less smoothed out so that they further enhance the view of sharp declination, toward the right and a gradual falling off, of this angle of declination progressively farther to the right. Also all curves start their initial descent at approximately the same point, namely, with scores lying between 7.9 and 9.24. This they had done in the unexpurgated curves of Graph *A*. Parallelness is most greatly, dis-

turbed, as between curves, at positions 2 and 3, and naturally at the right-hand terminals. All curves, considered together, show approximately the same declination in Graph *a* (after deletion) as in Graph *A* (before deletion, when the contamination of the short series and of the perfect recalls was assumed to be present).

The Point of Fixation in Relation to Left-to-Right Loss of Effectiveness of Letter-Positions

Other investigators [7] have suggested that when the subject is required to fixate a point to the right of the center of the exposure field, in range experiments, the results obtained turn out to be the reverse of those which we have presented. It is argued, moreover, that if the subject is allowed freedom of fixation, by the instructions given and special technique adopted, he will, in most instances, due to habit, fixate the left-hand side of the exposure line of letters. In Miss Johnson's experiment [8] four devices were adopted to bring about central fixation. (1) Specially typed instructions urged that the fixation be directed at the center of the exposure field; (2) these instructions were specifically reiterated at the beginning of each experiment; (3) practice exercises at each experimental session were given by the presentation of four pre-experimental exposures to habituate the subject to central fixation; and (4) very short, average, and very long lists of stimuli, each printed in the center of the exposure card, certainly were not conducive of fixation very far from the center; the letter-lists, presented in a random order, otherwise would result in the subject's being caught frequently, during an exposure-period of only 150 σ, staring at blank cardboard instead of the letters of the alphabet presented. In view of the above considerations, the writer is convinced that in her experiment fixation of the center of the field was the rule rather than the exception. And, it is recalled, the conditions in her investigation as described produced composite Graph *DD*, similar in all essential points to Graphs *D* and *d* of the present study. Hence,

[7] At the University of California. I believe the investigations referred to have not, as yet, been published.

[8] *J. Applied Psychol.*, 1928, **12**, p. 87f.

one is tempted to assert that central-fixation was characteristic of the 131 subjects in this work. Assuredly, a shortened exposure-period of 100 σ was an instrument, along with the use of cards varying in their numbers of letters from four to nine, presented in random orders, to facilitate and make imperative fixation at the center of the field. Moreover, instructions in the later study also encouraged central fixation; but, since no one experiment in the later investigation lasted longer than two laboratory periods (most, if not all, being consummated in one), opportunities of making the central fixation habitual were lacking. If we were to apply to the results a statement by Huey,[9] namely, that if a context is *unknown* the items grasped will lie to the right of the fixation-point and that if it is *known* perception will apply largely to the left of the fixation-point, the fixations of our subjects were directed to the left of center, inasmuch as the context was wholly lacking in every presentation, the various letters being always in nonsense arrangement, and the highest efficiency of recall accumulating towards the left of each card. For this comparison, however, Huey ' statement is, at best, ambiguous. It is to be recalled that he obtained this result from only a few cases and in reading experiments rather than in range of attention techniques. The writer has for many years been very skeptical of this discovery; he went so far as to say, in 1924,[10] that several similar results cited by Huey were of doubtful validity, such for example as the statement that reading is more easily effected during an exposure of 10 σ than in a longer exposure, —and that when the context is unknown the reader's range is increased. The exceedingly great tentativness of these results is such that one hesitates to apply them in converse fashion to determine what was the predominating fixation of the 163 subjects who served in this and the 1928 investigations. In fact, Dodge[11] has demonstrated that no matter where the eye-fixation is supposed to be directed, during even

[9] *The Psychology and Pedagogy of Reading.* N. Y., Macmillan Co., 1908, Part I, 15–184; *cf.* p. 56.

[10] An Investigation of Proofreaders' Illusions. *U. of Oregon Pub.*, October, 1924, p. 147f.

[11] *Psychol. Rev.*, 1904, 11, 1–14.

short exposures there are unsteady, roving movements of the eyes and such movements do not follow fixed and regular patterns, either to the right or the left, up or down. In addition, Ruediger [12] has demonstrated that "even a rapid reader does not use all his retinal extent available for seeing words"; that reading-rate "correlates highly (+0.81) with the smallness of reading pauses per line"; and that "reading rate is in the main determined centrally, by the rapidity with which meaning is aroused after the words are seen." Similarly, Hamilton [13] has demonstrated that "certain portions only (of the retina) are selected, *viz.*, the focal area or the *area of distinct attention* and of the right-hand marginal impressions" constituting what is called 'the reading field'; that a "probable cause (of inequality of perceptual distinctions) is to be found . . . in the fading out of the impressions on the right-hand margin of the visual field . . . during the time required for the attention to pass from left to right over the separate . . . impressions"; that 'the extreme brevity of the period of (eye) movement, about 30 σ', affords an insufficient 'time for the phase of interpretation' which, hence, must occur during the later stages of the fixation-pause itself; and that "a third factor which influences the (inequality of perceptual distinctions) is the character of the process of attention as independent of the area of most distinct vision on the retina", this relative independence resulting in "a margin of impressions of which only those on the right-hand could play any immediate part in the normal reading process, owing to our peculiar method of reading from left to right." Finally, Chou,[14] employing a quadrant tachistoscope which guaranteed, by its shutter device, central exposure, if not also initial central fixation, has shown the preeminence of left-to-right reading in so far as the effectiveness of attaining meaning is concerned. In detail his results which are applicable are as follows: "From the results of 16 Chinese subjects, left and

[12] The Field of Distinct Vision. *Archs. of Psychol.*, 1907, 1, No. 5, 68 pp.; *cf.* p. 68.

[13] The Perceptual Factors in Reading. *Ibid.*, 1907, 1, No. 9, 56 pp.; *cf.* pp. 52 and 54.

[14] Reading and Legibility of Chinese Characters. *J. Exper. Psychol.*, 1930, 13, 332–351; *cf.* p. 351.

upper halves of Chinese *characters* [15] have been found more suggestive of the whole-characters than either right or lower halves, since they take shorter time to read and yield more reproduction of the supposed whole-characters." But "the fact that right and upper halves of the *shutter* and *line* give faster reading and yield more reproduction is perhaps due to primacy (of memory) and right-handedness." A large majority of our 131 subjects were also right-handed and right-eyed, a fact ascertained by later tests, and yet in spite of this we have obtained the left-to-right loss in effectiveness depicted in the various graphs. We must await opportunities of the future to settle beyond all question this matter of the influence of fixation on the order of effectiveness of letter-positions. It is our strong belief, however, that long years of habituation to left-to-right reading will produce results similar to ours, regardless of fixation and regardless of whether or not eye-movements are ruled out of consideration.[16]

Is the Left-to-Right Decline of Letter-Position Effectiveness Due to Perception and Attention or to Priority (Primacy) of Memory?

We must frankly admit that we obtained no data upon, and had no set instructions or other technique to control, the manner in which each subject would reproduce in writing the letters he recalled as having been perceived during the exposure. General observations of the subjects during their recalls suggested strongly that most of them did reproduce from the left side of the recording blank towards the right. It is not impossible that this fact is the major factor explaining our results, although the writer doubts it. This has always been a dubious question in the matter of the range of perceptual attention—how much are the results characteristic

[15] Italics and parentheses are mine.

[16] In fact, Dallenbach (*J. Exper. Psychol.*, 1920, 3, pp. 206–207), working with three subjects, including himself, discovered that in only one of his subjects did fixation appear 'to be the sole condition necessary for clearness'. Another of them revealed other conditions, than fixation, as affecting clearness. Dallenbach finally decided that the first subject alluded to possessed other conditions working in the same direction as fixation so that fixation was not the sole determiner of clearness.

of perception and how much of memory? Our grasp of the literature affords no description of a technique which has been effective in solving it.[17] A future investigation is being plan-

[17] Our study of 1928 has been criticized by Dallenbach (*Psychol. Bull.*, July, 1930, p. 501) in the following words: "Crosland and Johnson's study on range of apprehension is *mistitled.* (Italics mine.) The authors are not interested in the question of range (they do not even compute constants for their *S*s), but in determining how legibility of printer's type is affected by inter-letter hair-spacing and by the characteristics of individual letters—a topic still further removed from the scope of this review. We must note, however, their use of the technique of the *traditional range-of-attention experiment*, and their *assumption* of an invariable relation between report and legibility —an assumption disproved in range-of-attention experiments of the past decade. The conclusions regarding legibility are, therefore, unwarranted inasmuch as they transcend the results," etc. In view of this devastating summary it may be surmized that the above sentence will bring forth from him the allegation that we have failed to keep abreast of the literature or, at any rate, we have failed to encompass *his* technique. In fact one is impressed with the almost wholly negative attitude—couched in language which is, to say the least, iconoclastic—which he has displayed toward the work of Ford, Sen Gupta, Rodrigue, Wolf, Crosland and Johnson, Gundlach and Rothschild and Young, Wilcocks, Meisenheimer, Wheeler, and Prengowski. Such an extensive list of curt criticisms, to say nothing of his rebuttals of Morgan, Wever, H. M. Johnson, Oberly, and Fernberger, causes one to wonder what is that technique of the 'past decade' which can so successfully supplant the 'traditional range-of-attention experiment' and what technique is it that can distinguish between attention of perception and the report of memory. Lest his charge against Talhouët (*Psychol. Bull.*, 1928, p. 496), that the latter was discussing attention in terms derived from 'a generation ago', be added to that quoted above a charge from which H. M. Johnson failed to escape (*Psychol. Bull.*, 1928, p. 504), I wish to say that I have gone to the pains (substantial, to say the least) and labor of reviewing his various controversies with Morgan, Wever, H. M. Johnson, Oberly, and Fernberger; of rereading his three summaries of the literature on attention (*Psychol. Bull.*, 1926, 1928, and 1930); and of reperusing the various papers (approximately thirty in number) of his and his students, such as the papers of Cassel, Gill, Friedline, Glanville, Dewey, Guilford, Burke, and Meads. In addition I have reviewed the papers of Geissler (1909), Archer (1916), Woodrow (1914, 16), McQueen (1917), Curtis and Foster (1917), Bollman (1920), Newhall (1921), McComas (1922), Cooper (1928), Verwoerd (1928), Wilcocks (1928), Philip (1928), and Ford (1929). All of the above literature may be condensed for the purposes of this discussion, I believe, into the following summary: (1) Dallenbach believes that he knows what attention is and what are its various characteristics and just how to measure it and them, as well as believing that he is lacking in that 'provinicialness' which he has charged (*Psychol. Bull.*, 1928, p. 494) Henning as possessing. (2) 'Range' is a term which should never be applied to 'attention'. (3) 'Range of attention', if any, is always unity. (4) Wirth's and Geissler's 'many-level' type of attention does not exist. (5) Many factors determine a subject's report other than the factor of mere clearness of perceptual attention. (6) Clearness is of two sorts, attributive and cognitive (the latter, I believe, contaminating report in company with other factors), to make no mention of objective and physiological clearness. (7) Most, if not all, techniques, including the 'traditional' range-technique, the reaction-time and the efficiencyof performance

ned with a technique which may solve the question for this particular type of project. It is wise to await the outcome of the undertaking before hazarding guesses. Yet the fact remains that in two experiments, separated by two years in time, with different groups of subjects, 163 in all, with some-

techniques, have proven themselves failures as measures of attention, particularly of range of attention. (8) One technique, at least, which can stand up under fire, has been perfected in the 'last decade', and that is the technique of Dallenbach and Cassell (of which several of us have been so woefully ignorant or to which a few of us have been so callous). (9) This technique involves, above all else, (*a*) introspection, of a very refined degree, as to the number of items which stand out, in a perception, at any moment of tachistocopic exposure, as well as to the various levels of such 'standing-out' and the degrees of assurance with which the subject 'reports' numbers of items and levels of clarity; it involves (*b*) the taking of reaction-times and the averaging of them for each level of clarity and then the running of correlations between the average time and the average clarity of a given class-interval (a procedure doubly condemned by statisticians); (*c*) the application of a method of obtaining constants; (*d*) the employment of 2 or 3 subjects; and (*e*) the giving of each subject 3 to 4 months' practice in introspection.

In summary, he has done, with the literature of the 'last decade', exactly what he excoriates Wheeler with doing (*Psychol. Bull.*, 1930, p. 508): the literature of the "last decade" "is entirely of his own making." It by no means supports his concepts without limitations and exceptions, and his technological methods have not won such universal acceptance as to color and permeate a majority of the literature of this more recent period. Therefore, a result, or a technique, or a criticism which does not harmonize with his concepts is merely 'traditional' or has arisen out of ignorance of the literature. It is my candid belief that one who uses only 3 subjects, runs correlations between two sets of data from only 3 subjects, and correlates class-interval means instead of parallel single-item scores is in no position to challenge all other investigators.

With regard to my own deviations from his prescribed technique, I may state that, while I was reared in a tradition of introspection, I no longer expect to follow the introspective methodology when and where I believe I can obtain reliable and valid results in lieu of it. I have given up introspection very largely because I have discovered that it is more liable to error than are objective methods. It is impossible to prevent introspections from being greatly influenced by those very biases of which hint was made in the preceding paragraph, nor are introspective results easily quantifiable or checkable against the results obtained by other investigators. Dallenbach is quite right in saying that Miss Johnson and I had no interest in the 'question of range', if 'question of range' presupposes what apparently it does for him; nor were we interested in 'pure' experimental work (*Psych. Bull.*, 1928, p. 510), again if to be 'pure' requires his concepts and his technique. Nor, I may add, will our interest ever grow or attach itself to his line of 'pure' approach, if to do so it must be fed upon what I regard as the futile or near-futile controversies between himself and each of Morgan, Wever, Johnson, Oberly, and Fernberger.—It seems to many people exceedingly disappointing, if not distressing, and at best a most curious situation, that most psychologists should be wrong and only one right and that only one should be conversant with the literature and with the techniques of the 'last decade'. Watson did not win approval and support in all quarters when he inveighed against the 'magic' and the secret 'initiation

what different apparatus and techniques, the same result has been brought to light: namely, recorded recalls of letters of the alphabet, supposedly perceived just previously, do

into the occult' of those few who can give perfect introspections. But a perusal of this mass of literature, to discover who holds the secret and that the secret is the introspective wand of a special coterie of interpreters, all else being charged to ignorance or to the lack of scientific interest, is enough to drive many into Watson's camp.

I have discovered nowhere in the above literature a method with which one can distinguish between what was given by attention and what was given by report in its own right, between what was perceived and what is recalled as having been perceived. This problem, and its discovery, is not new; nor has the solution, as far as I can find, been effected, in the 'last decade' or any other decade. Therefore, if we seemed to have assumed that attention and report are co-variables, we did so because of the lack of a better technique; we believe that we are no less deserving of generosity in this matter than are other investigators, past and present. Certainly no such assumption was predominant in our minds throughout the experiment and its subsequent publication. We tried out the technique, got results which pass statistical muster, and specifically stated that we had no intention or inclination to enter the controversy concerning range or concerning range, cognition, and apprehension. We are willing to let the results of 1928 and the present investigation stand for themselves. But let us suppose that such a technique requires refined and detailed introspections as to levels of clearness and the number of letters of the alphabet contained in each level. Will any introspective 'report' be more accurate and reliable and valid than the so-called objective, written report of the subject of the letters he recalls from the cards which were flashed before his eyes? Are we to suppose that many factors determine 'report' but only one factor, namely the number of letters in a given attention level during perception, determines introspective report, expression of certainty, or what not? Did not Miss Washburn, in her presidential address, show that introspection is itself a form of report-reaction? And has not the chief fault of introspection always been that it is very easily contaminated by many and varied factors? Why should the subject's writing down of the letters of the alphabet as recalled be any more untrustworthy or more influenced by memory, etc., than his oral or written statement that "five letters stood out with equal clearness of a focal sort, and two letters were barely glimpsed, they being so marginal," so that two levels of clarity were present? Are we to suppose that introspection on the spot is not affected by memory or imagination? or that the retailing of the introspection to the experimenter or upon paper a moment subsequently is independent of memory or imagination? Or is the situation greatly improved when the subject is asked to grade his objective recalls by introspection of the degrees of assurance with which he has recalled? I believe it is notoriously true in the literature that the correlation between degrees of certainty and the objective correctness of recalls is very low. Elsewhere (*Psychol. Monogs.*, 1921, No. 130, pp. 125–133) I have tried to develop the thesis based upon introspective facts, that degrees of certainty, feelings of certainty, are themselves subjective constructs and are just as imaginatively and conceptually developed as are fullgrown hallucinations. Koffka, Helen Clark, Perky, Martin, Ogden, Henderson, Philippe, and Kuhlmann (*Psychol. Monogs.*, 1921, No. 130, pp. 10–29) developed a somewhat similar thesis, as did also Bentley and Whipple.

If there is a technique in existence, other than that alluded to, by which one can effectually differentiate between attention or perception and the report thereof, the writer would appreciate an opportunity of overcoming his ignorance of the literature

possess a sharply and then more slowly declining effectiveness from left-to-right by letter positions.

Summary

1. The composite graphs, of 1928 and of the present investigation, constructed of all the scores from short, average, and long series of letters without distinction between series but arranged according to letter-positions from left-to-right, reveal a rapidly declining and then a more slowly descending effectiveness of letter-positions from left to right, up to the 6th letter-position, after which there is no further decline.

2. Three composite graphs, *DD*, *D*, and *d*,—the first, of 1928, constructed from the records of 32 subjects,—the second constructed in this investigation from the records of 125 subjects,—and the last constructed from the records of *D* from which have been deleted the records concerned in all perfect reproductions of cards,—all possess most striking similarities.

3. The records contained in Graph *d*, when treated separately by series, thus to distinguish the series of four, five, six, seven, eight, and nine letters each (for the first trial), as in Graph *a*, reveal practically no differences between series and make obvious the general fact of sharply declining and then more gradually declining effectiveness of recall by letter-position from left to right.

4. The general tendency, pronounced and unmistakable, of the left-to-right decrease of letter-position effectiveness, referred to in connection with Graphs *DD*, *D*, *d*, and *a*, was *not* obscured in the 'contaminated' records pictured in Graphs *A*, *B*, and *C*, records containing all scores, even those emanating from the shortest series (the 4's) and from cards characterized by perfect reproductions.

5. The composite method of graphing had its most serious effect in obscuring the attention-value of the final and near-final characters to the right of the individual series. Every series, without exception, in the individual series graphs, depicts this augmented attention-value at the extreme right. A much smaller effect, resulting from the pooling of series

results, was the bulge in the composite graphs, or increased effectiveness, in letter-positions 3 and 4 undoubtedly resulting from the scores of Series 4 and 5. Strange as it seems, positions 1 and 2 were little if at all affected by the composite method of graphing or by the 'contaminating' influences of the records from cards recalled with 100 percent efficiency.

6. The deletion, from Graph *A*, of all scores from cards which had been perfectly recalled, as depicted in Graph *a*, resulted in the greatest reduction in the efficiency of recall for letter-positions 3 and 4, and affected more seriously Series 4 and 5 than any or all other series. All series considered together were affected to the degree of a reduction of only 4.3 percent. Positions 2 and 5 were slightly affected, to an amount equal to that for all series. The near-parallelness of the curves of a declining left-to-right effectiveness for all series is very striking in Fig. *a*. All curves start their descent at approximately the same point, regardless of the deletion and regardless of the shorter series employed along with the longer; —that is, position no. 1 is high in effectiveness independently of all such 'contaminating' influences. Likewise, no method of treating the results (save that of pooling) and no arrangement of longer series are of a sort as to eliminate the higher attention or recall values of right-hand terminal letters.

7. Preliminary expectations (and data from another investigator's work) [18] have received little or no support when all perfectly-recalled cards were excluded from the data;— Series 4 contained only 52 percent of cards recalled perfectly and Series 5 only 20 percent. All other series contained approximately zero percent of perfectly recalled cards. Nor were the remaining cards, in Series 4 particularly, free from the left-to-right sharp declination of effectiveness by letter-position. Graph *E* reveals as a matter of mere coincidence that the percentages of cards recalled without error parallel the loss of effectiveness of letter-positions from left to right.

[18] M. A. Tinker, How Formulæ Are Read. *Amer. Jour. Psychol.*, 1928, 476-483; and *Psychol. Bull.*, 1929, **24**, p. 612.

Conclusions

1. There were, in general, three forces working to produce the foregoing results: (*a*) a very strong tendency, unmistakable and ineradicable, of a sharply and then gradually declining loss of effectiveness of left-to-right letter-positions in recall; (*b*) a much weaker tendency for the perfect recalls, in the shortest two series, to raise the recall values of letter-positions three and four; and (*c*) a tendency for the final character or letter-position to possess a higher attention or recall value than its left-hand antecedent, a tendency which, however, was much weaker than that of the general declination of (*a*), above.

2. The composite method of graphing was harmful in that it obscured tendency (*c*). It also, with less serious effects, somewhat obscured the differences existing between Series 4 and 5 and all other series.

3. Conclusion 2, above, suggests that the questions involved were real experimental questions and that they could not be adequately and satisfactorily answered by *a priori* or rationalistic methods, nor could they be answered with data from a very small population of subjects.

4. This investigation has proved to the author that it is highly desirable and wise to employ in investigations of this sort series lying both above and below the average subject's range and that these series should be presented in chance orders to the subjects. The literature dealing with range of attention and with thresholds in general bear out this contention, of the desirability of random arrangements of below-range, range, and above-range numbers or quantities of stimuli. Again it is apparent to us that the question of the effects of long and short series and of their random orders of presentation is entirely an experimental question, one not susceptible of a deductive type of answer nor of an answer based upon data from merely a few cases. Certainly an experiment to compare the two techniques, of obtaining merely percentages of cards recalled with full correctness and of obtaining individual series scores, letter-position by letter-position, both before and after expurgation of perfect recalls, appears to be a justifiable and reasonable pursuit.

5. We believe that left-visual-field fixations cannot be appealed to to explain our results, in view of Miss Johnson's earlier techniques, the techniques in this investigation, and the great similarities between the two sets of results.

6. The influence of primacy of memory, that is, of left-to-right manners of manual reproduction by right-handed subjects cannot dogmatically be rejected as a partial cause of the left-to-right decline of letter-position effectiveness. This again is an experimental problem which it is hoped will in the near future be satisfactorily solved.

(Manuscript received September 5, 1930)

A FURTHER STUDY OF CONFIGURATIONAL LEARNING IN THE GOLDFISH[1]

BY F. THEODORE PERKINS

University of Kansas

Introduction

The field of problem-solving in animals lower in the evolutionary scale than the mammal and bird has been little exploited. However, the value of such research cannot be overestimated, for the human subject probably has emergent behavior-patterns in his early development similar to those of the lower animal.

The fish offers fertile material for such investigation, not only because of its sensory capacities (acute vision in particular) but because of its adaptability to laboratory conditions, its motility, and ease in rearing. The goldfish (*Carassius auratus*) readily satisfies the above requirements. Perkins and Wheeler[2] have shown that this animal has striking insight into visual problem-situations. This study demonstrated that goldfish were not only able to choose their food from three absolute light-intensities merely changed in position, but were also able to choose a light in relation to two others, when the absolute intensities and positions were changed after each response, indicating that the responses were not made to an isolated stimulus, but rather to a total stimulus-pattern. Transpositions both up and down the intensity scale, in the case of the fish that responded to three absolute intensities, failed appreciably to reduce their degree of accuracy—further evidence for a pattern response. A continuous emergence of new insight into the problem, the result of a gradual maturation of the behavior pattern, was

[1] The writer wishes to express his utmost appreciation to Dr. R. H. Wheeler, whose valuable suggestions both during the experimentation and in the writing of the results made this study possible.

[2] Perkins, F. T., and Wheeler, R. H., Configurational Learning in the Goldfish. *Comp. Psych. Mono.*, 1930, 7, No. 1, 1–50.

clearly demonstrated. This study, however, raised certain problems that deserve further experimental consideration. (1) There was no evidence to prove whether the responses to the light situations were made on the basis of brightness or of color, as the light intensity was controlled by means of rheostats. With a lowering of intensity concommittantly came a change in yellowness. (2) The animals seemed to prefer the medium light, regardless of intensity, a fact which needed further substantiation. (3) The conditions for handling the fish were not well controlled so that certain of the animals died. (4) Training the animals was continued after the transpositions, where incorrect responses were made, thus according to Warden,[3] possibly forcing the animals to make relative or pattern-judgments, whereas before they had not. Although Perkins and Wheeler's learning curves directly disproved this supposition, Warden's hypothesis was recognized as a worthwhile precaution. (5) Further quantitative and qualitative facts regarding the behavior of the goldfish in problem-situations fastens another link in the chain of comparative evidence for the genetic development of insight. (6) A detailed comparison of the behavior of the goldfish when presented with a constantly changing pattern of lights, both in intensity and position, as contrasted with the behavior of those presented with a pattern changing only in position, was desirable. The present study was undertaken in an effort to solve the foregoing problems.

Apparatus

The apparatus used in this experiment (see Fig. 1) consisted of three food-compartments (10 cm by 10 cm by 10.5 cm) [4] arranged in a semicircle with doors (5.5 cm by 5.5 cm) in the center and at the bottom of the compartment. These doors opened into a choice-compartment (33 cm by 15.5 cm by 10.5 cm) which was large enough so that lines drawn perpendicular to each compartment door joined at the rear of the

[3] Warden, C. J., and Rowley, J. B., The Discrimination of Absolute *Versus* Relative Brightness in the Ring Dove, *Turtur Risorius*. *Jour. Comp. Psych.*, 1929, **9**, 317–338.

[4] The order of dimensions will in all cases be length, width and height.

choice-compartment. This enabled the fish to obtain practically the same perspective of each compartment door, when in the center of the choice-compartment. The backs of the food-compartments were plates of milk glass, two in each compartment (2.5 cm apart) through which the light source passed. The sides of the compartments were extended so

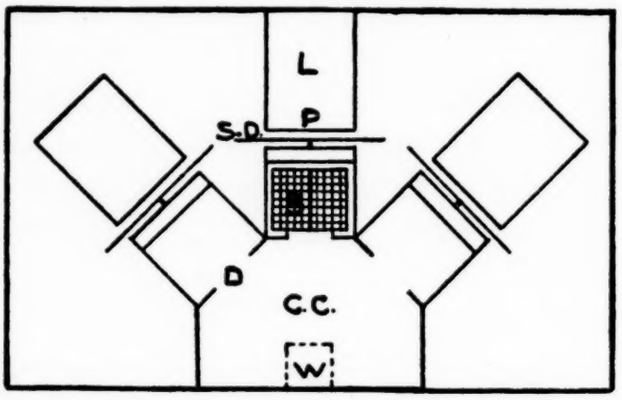


Fig. 1

C. C., choice-compartment.
D., doorway to food-compartment.
L., light-source compartment.
P., milk glass partition.
S.D., stimulus disk.
S., excluding screen.

that another compartment was formed (the light-source compartment; 15 cm by 10 cm by 10.5 cm) large enough to hold a socket and a 60 w. lamp. The light socket was passed through the top of the light-source compartment and the bulb extended downward. The light-source compartment was sealed except on the front side where there was another plate of milk glass. Thus the light passed through three plates of milk glass before it entered the food-compartment. Between the food-compartment and the light-source compartment was left a narrow space, into which there tightly fitted a metal disk (21 cm in diameter) that pivoted from the center of the top wall on the food-compartment. In each of the three disks was cut a series of five apertures spaced in such a way that when one of them was centered upon the light source, light would penetrate through only that one aperture. The holes ranged, in size, as follows: .04 cm, .5 cm, 1.0 cm,

2.2 cm, 4.4 cm. This meant that by revolving any one of these disks 360 degrees five different intensities of light could be obtained without changing the light source. This eliminated the yellowness that occurred when rheostats were used. On the inside floor of the food-compartment, at the front edge, there was hinged a course meshed screen the same size as the floor of the compartment. A black string was attached to the free side of the screen and passed through the top of the food-compartment, over the choice-compartment, and through holes made in a wooden screen, between the observer and the apparatus. When the observer pulled on the string the screen lifted, excluding the animal from the food-compartment, without injuring it. An aperture was cut in the wooden screen, between the observer and the apparatus, to permit observation of the animal in the choice-compartment. The entire apparatus, which was situated in a darkened room, was placed in a large aquarium and submerged in water to a depth of about five inches. A square wire cage (6 cm by 6 cm) open at the top and bottom was placed on the floor of the choice-compartment at a point equidistant from the door of each food-compartment. The floor of the choice-compartment was painted black.

A block of square wire compartments, of the same mesh as the wire cage and the excluding screen, open at the top, was submerged in the other end of the aquarium to the same depth as the experimental apparatus. The fish were kept in these compartments at all times except when actual experiments were being performed.

In apparent foot-candles, determined from photometric readings taken with a McBeth Illuminometer, the five light intensities were as follows, with the largest aperture numbered 1:

TABLE I

LIGHT VALUES

No. of Aperture	Apparent Foot-candles	Diameter
1	10.55	4.4 cm
2	4.8	2.2 cm
3	.73097	1.0 cm
4	.21372	.5 cm
5	.01897	.05 cm
0*	0	0

* Zero was added in some cases by turning out the light in a particular compartment.

PROCEDURE

General Method: Care of Animals.—As the same general procedure was used in all the groups it will be described here and will be understood to apply to all groups, except one, the control group. The fish were fed in the experimental apparatus and never in the retaining cages. Each fish was kept in an individual cage, the number of which determined the animal's number and its position in the group. The animals were practically all the same size, and as near as could be determined, from the same school. Dippers were used in transferring the fish from the retaining cages to the experimental apparatus. All the animals in the experimental group were run on the same days, thus eliminating any differences between sub-groups due to different time intervals.

Experimental Method.—Turning to the method used in the actual experimentation, the animal was first transferred from the retaining cage, by means of the dipper, to the wire pen in the center of the choice-compartment. (The subjects were experimented on separately.) Food of a commercial variety was placed, in the same quantity, in each of the three food-compartments in order to eliminate any possibility of a response on the basis of olfactory cues. Control of temperature was unnecessary because the lights in the food-compartments were of equal intensity, and thus, if at all, would heat each compartment to the same degree. The problem-situation was then arranged by means of the metal disks, and the wire cage raised from the animal. In all cases the problem-situation was arranged before the animal was released, so that it could obtain an equal perspective of the problem at the outset. If the animal entered the correct compartment (determined by the group of which it was a member) it was allowed to remain there about 20 seconds to eat, and if at the end of this period it had not returned to the choice-compartment, the experimenter tapped on the top of the food-compartment. This always aroused the animal and let it return to the choice-compartment of its own accord. If, on the other hand, the animal entered the wrong compartment, the experimenter immediately pulled the string attached to the

excluding screen in that compartment, whereupon the animal quickly left the box. Then the experimenter gently placed the wire cage over the animal again and guided it to the same position as for the first response. The next problem-situation was then arranged and the procedure repeated. Each fish was kept in the apparatus each day until it had made five responses or, in case it had not made five, for a period of ten minutes. It was decided arbitrarily to record a response when the fish went into a compartment as far as the rear fins. The problem was considered learned when the animal could make ten consecutive correct responses out of ten trials. There were in all 48 subjects that made a total of 3,812 responses.

Procedure for Individual Groups: Experimental Group.— The experimental group was divided in two parts, which we shall call the dynamic and static groups, each comprised of nine fish. *Dynamic Group:* By the dynamic group we mean that a constantly changing stimulus-pattern was presented to the subjects, that is, after each response of the fish there was a change made not only in the relative position of the three lights, but also in their absolute intensities. This meant that there was a transposition after each response, with the relationship of brightest, middle and dimmest remaining throughout. Table II shows the ten stimulus-patterns used

TABLE II

STIMULUS-PATTERNS FOR DYNAMIC GROUP

Combination	Left	Center	Right
A	3	1	5
B	4	2	1
C	5	1	4
D	2	5	3
E	3	1	4
F	4	5	2
G	1	3	2
H	3	2	4
I	5	4	3
J	2	1	5

with this group. No one pattern occurred more than once in any day's trial, and on each day a different order of the patterns was used, by starting at a different point in the series.

Only five out of the ten patterns were employed in one day. The dynamic group was divided into three sub-groups.

Sub-group A contained three fish, no.'s 1, 2 and 3, which were trained to secure their food from the brightest light out of the three presented, regardless of its position or absolute intensity. Sub-group B was composed of fish no.'s 4, 5 and 6 trained to respond to the medium light as a food-goal. Fish no.'s 7, 8 and 9 were the members of Sub-group C, which were trained to respond to the dimmest light in any pattern.

On the last day the members of each of the sub-groups were given a test containing ten strange and novel stimulus-patterns (see Table III). It is important to note that the training was discontinued at this time and more important still, to note, in the light of Warden's objection, that the fish needed no further training. (Training means punishment.)

Static Group.—The other half of the experimental group was designated as the static group. By this is meant that the fish were presented with three absolute light intensities, the brightest, the medium and dimmest of the series (1, 3, 5) merely changed in position after each trial. After the animals had learned the problem they were transposed both up and down the intensity scale, and again the training was discontinued after the transposition. This group was also divided into three sub-groups.

Sub-group D included fish no.'s 10, 11 and 12, selected to choose the bright light (no. 1) regardless of position. After this problem had been learned the animals were transposed to a combination of lights much reduced in intensity, namely, 4, 5 and 0. When they had made 100 percent correct choices a transposition to the combination, 2, 3 and 4 was introduced. After perfect performances were again reached, the animals were given a rigid test, in which ten stimulus-patterns appeared with the various light intensities in novel spacial and relationship combinations. This was the same test given to Sub-group A of the dynamic group. In no case was the training of the animals carried on after the first transposition had been made. This fact holds true, as well, for the sub-groups that follow.

Fish no.'s 13, 14 and 15 made up Sub-group E of the static group. These animals were trained to choose the medium light (no. 3). When the problem was mastered the animals were transposed down the intensity scale to lights 4, 5 and 0 and at the same time the training was discontinued. A pattern of lights 1, 2 and 3 was presented after the foregoing problem was learned. A test, the same as that for Sub-group B of the dynamic group, was presented after the 1, 2 and 3 situation was learned in which the lights appeared in the novel arrangements already mentioned.

Sub-group F was composed of fish no.'s 16, 17 and 18, which were trained to choose the dim light (no. 5). After this problem was learned, the animals were transposed to the combination, 4, 5 and 0, at which point the training was discontinued. Another transposition was made when this problem was adequately learned, to the combination, 1, 2 and 3. No test was given at the end of the experiment.

Control Group.—In order that some basis might be found for checking the performances of the experimental group, a control experiment was carried out. Thirty fish were used, divided into two main groups and six sub-groups. In this case the animals were not kept in separate compartments, except as to sub-groups, for the reason that the animal remained in the experimental apparatus until the experiment was finished. The fish were placed individually in the wire cage in the center of the choice-compartment and the stimulus-pattern of three lights was arranged as before. The animal was then released and allowed to make a response. However, this time no punishment was given, although the food remained a constant factor in all the compartments. In other words, the animal made a free choice, which means that its light preference was being tested. Each animal made 20 consecutive responses, with the lights in all cases changed in position after each response (the animal as before starting from the wire cage to make a response) so that, except for the training, the conditions were the same as before. The control group was divided into two general groups, the dynamic and the static controls.

The *dynamic control group* was composed of five fish, which were given the stimulus-patterns used in the dynamic experimental group (see Table I). That is, the combinations were changed in absolute intensity as well as in position after each response of the fish.

The *static control group* was composed of 25 fish, presented with only three light intensities, merely changed in position. This group was divided into five sub-groups, each of which contained five fish. The fish in Sub-group G were presented with the lights 1, 3, 5; Sub-group H with 2, 4, 0; Sub-group I, 1, 2, 3; Sub-group J, 2, 3, 4; and Sub-group K with 3, 4, 5. It will be noted that in sub-groups G and H the intervals between the lights were greater, consisting of one step, while in sub-groups I, J and K, the intervals were smaller with no steps between the lights, that is, the lights were adjacent in intensity.

Results

Experimental Group.—The curve of fish no. 1's performance (see Fig. 2)[48] showed no particular learning until after the seventh day, at which point a steep climb began. The curve then fluctuated up and down until the 40th day when a plateau was reached. This plateau lasted for six days and presented evidence for the setting in of an irradiation pattern. It represented, probably, a period during which maturation was taking place, for at the conclusion of the plateau the animal's curve quickly ascended to 100 percent and remained high to the end of the experiment. The test then given the fish failed to result in any increase in errors (see Table III). The various steep ascents in the curve seem to indicate that the solution of the problem was not a gradual affair, but quite sudden—evidence for the principle of insight, not stamping in by trial and success. The fact that the animal responded adequately to the strange patterns presented at the end of the experiment is further evidence that the learning was not a stereotyped affair, but one that was dynamic and flexible in character. If we tabulate the errors made to the medium and dim lights we may note some interesting facts. The number of errors to the dim light was very high at the outset and decreased at the same rate with which the correct responses increased, but the number of errors to the medium light was rather small at the outset and remained practually the same throughout the experiment.

The performance of subject no. 2 was very similar to that of no. 1 (see Fig. 3). This curve also ascended very rapidly at the end of the experiment, again evidence for the rapid mastery of the problem as soon as the animal had matured with respect to that problem. The test at the end of the experiment again failed to reduce the animal's degree of accuracy (see Table III). The errors to the medium and dim lights were distributed similarly to those of the preceding subject.

[48] Units along the abscissæ in all figures indicate numbers of days practice. Units along the ordinates indicate the percents of correct choices. Numbers beneath dotted lines represent light intensities as per Table I.

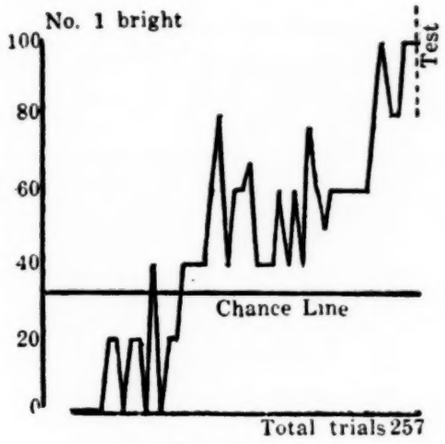


Fig. 2

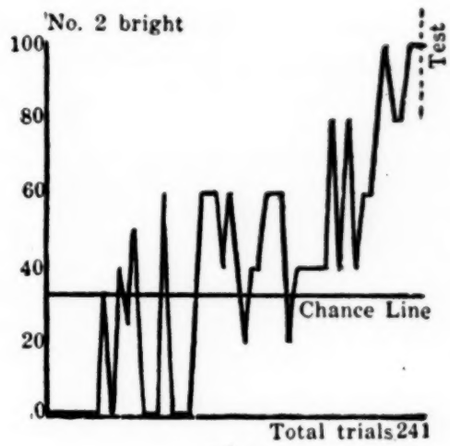


Fig. 3

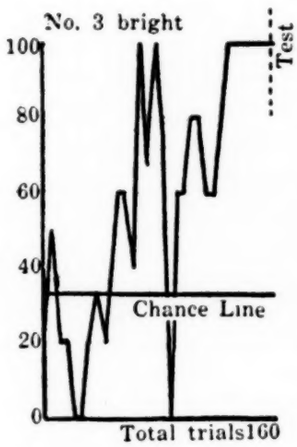


Fig. 4

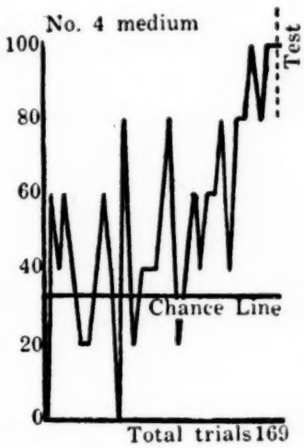


Fig. 5

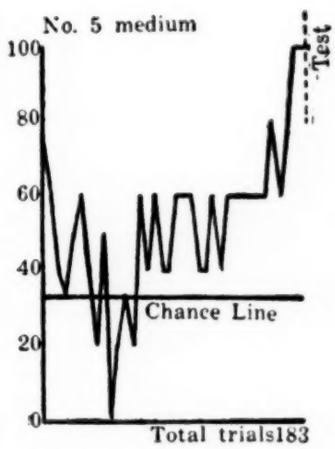


Fig. 6

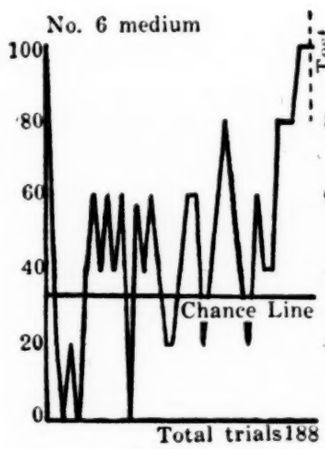


Fig. 7

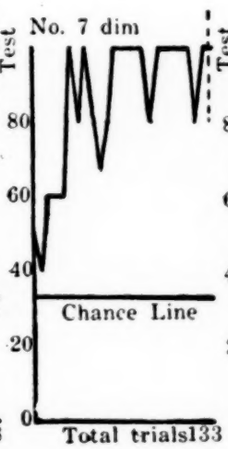


Fig. 8

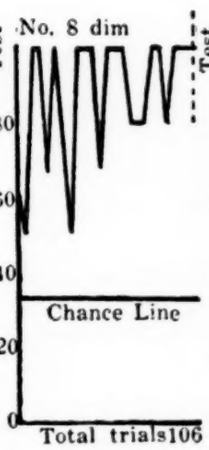


Fig. 9.

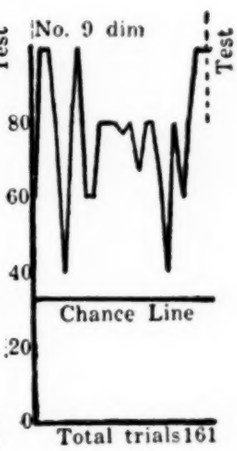


Fig. 10

TABLE III

TEST RESPONSES OF EXPERIMENTAL GROUP

Dynamic Group	Combinations									
Fish No.	225	313	444	550	121	111	433	515	333	123
1 bright...	2C	1	no	no	1L	1L	3R	1	3L	1
2 bright...	2L	1	C	no	1L	1C	3C	1	3R	1
3 bright...	2C	1	no	no	1R	1R	3C	1	3C	1
Fish No.	450	021	444	145	121	111	342	105	333	123
4 medium..	5	2	4L	4	2	no	3	5	3C	2
5 medium..	5	2	4C	4	no	no	3	5	3L	2
6 medium..	5	2	4L	4	no	no	3	5	3L	2
Fish No.	552	131	444	550	212	111	433	155	555	333
7 dim.....	5L	3	4L	o	2R	no	4	5R	5L	no
8 dim.....	5L	3	4C	o	2L	1L	4	5R	5C	no
9 dim.....	5L	3	no	o	2L	no	4	5R	5C	3R
Static Group	**Combinations**									
Fish No.	225	313	444	550	121	111	433	155	333	123
10 bright..	2C	1	no	no	1R	1C	3R	1	3L	1
11 bright..	2L	1	4R	5C	1L	1R	3R	1	3C	1
12 bright..	2L	1	4L	no	1L	1L	3R	1	3R	1
Fish No.	450	021	444	145	121	111	342	105	333	123
13 medium.	5	2	4R	4	no	no	3	5	3C	2
14 medium.	5	2	4R	4	2	no	3	5	3C	2
15 medium.	5	2	4L	4	no	no	3	5	3R	2

Legend: L, left.
C, center.
R, right.

Note: In the combinations the numbers stand for the light values, the order in each case representing positions left, center and right. The numbers opposite the subject-number indicate its response to the combination heading that column.

Fish no. 3, also responding to the brightest light in any three presented, exhibited a much shorter learning curve than either of the two preceding. It is characterized by steep ascents and one steep decline. These steep ascents again give evidence for the rapid growth of insight into the problem. The test situation did not lower the curve from the 100 percent mark. The error distribution in this case is not as clear cut, but there is some indication that responses to the dim are more numerous than those to the medium light (see Fig. 4).

Fish no. 4 responded to the medium light out of any three presented. The curve of its performance (see Fig. 5) is marked by steep ascents and declines, noteworthy

among which is the rapid ascent at the end of the experiment. The final test failed to increase its errors, the animal maintaining 100 percent throughout (see Table III). It is interesting to note that the animal made only one response to the bright light.

The curve of fish no. 5 (see Fig. 6), which also responded to the medium light. started relatively high, at 75 percent. The curve descended to a low mark on the tenth day, from which point it mounted to a level between 40 and 60 percent. Then there was a plateau at 60 percent, of six days' duration. This constant level again represents the maximum insight of the animal into the problem presented, and at the end of this period, when more maturation had taken place, the curve mounted rapidly to 100 percent. At the end of the experiment a rigid test was given, to which the animal responded with a perfect performance (see Table III). The errors show that there were few responses to the bright light, making the errors to the dim practically the inverse of the medium curve.

The curve of no. 6 (see Fig. 7), also trained to the medium light, commenced at 100 percent, probably for chance reasons. The record declined to low points on the third and fifth days, then mounted to a level between 40 and 60 percent, which was maintained, except for the fall on the twelfth day, to a period near the end of the experiment. At the end of this plateau the animal's curve mounted very rapidly to 100 percent, a level that was not broken by the test-situation at the end (see Table III). In this subject's errors there were relatively few responses to the bright light, so that the dim errors run opposite to the medium curve.

Fish no. 7, which was trained to choose the dimmest light of any three presented, gave a learning curve (see Fig. 8) that began at 50 percent and ascended very rapidly to 100 percent on the sixth day. Except for one appreciable descent on the tenth day the animal maintained a very high degree of accuracy to the end of the experiment. It will be noted that considerable 'over training' was carried out with this subject. The test with novel situations, given at the end of the experiment, failed to reduce the animal's accuracy (see Table III). If the errors of this subject are examined it can be seen that the mistakes to the bright started at 50 percent and then decreased rapidly to zero on the fifth day. Errors to the medium were few at the outset and ran reciprocally to the dim curve. This animal made very few errors throughout the entire experiment, and in this, we have demonstrated very clearly the rapid development of insight into the problem presented.

The curve of no. 8 going to the dim light (see Fig. 9) was similar to that of no. 7 in that it maintained a high level throughout the experiment. The curve started at 60 percent and mounted rapidly to 100 percent. The curve then had steep ascents and declines, but never went below 50 percent. The declines, however, decreased in extent as the experiment progressed. The test at the end was mastered correctly by the animal (see Table III). A consideration of the errors shows that there were *no responses* to the *bright light*. Errors to the medium, therefore, were the converse of the dim curve.

Fish no. 9, which responded to the dim light, started at 60 percent and ascended to 100 percent on the following day. After remaining at this level for two days the curve dropped, then mounted rapidly again and declined to a level of about 60 percent. Again the curve (see Fig. 10) mounted rapidly as the experiment was drawing to a close, and the test at the end failed to increase the subject's errors (see Table III). The errors show in this case a number of incorrect responses to the bright light, but the dim curve shows a preference for the dim light.

Fish no. 10, the first member of the static group, responded to the brightest light

of the three absolute intensities, which were presented in varying positions. The learning curve (see Fig. 11) of this animal started at zero and made several steep ascents in the first part of the experiment, but each time the curve descended again to the zero mark. On the 35th day the animal was not making more than 60 percent correct choices, but at this point the curve *ascended in one day to 100 percent* and after a slight decline maintained this level to the end of the experiment. At the end of the 42nd day the animal was transposed to the 4, 5 and 0 problem, and at the same time the training was discontinued. The animal made one error and then a perfect performance. After the 44th day the animal was transposed to lights 2, 3 and 4, which failed to lower the learning curve. On the last day the rigid test that was given Subgroup A was presented and resulted in a perfect performance (see Table III). The errors show that responses to the medium light started at 60 percent, and higher at the first of the experiment, and decreased toward the end. The errors to the dim were few at first, reached a peak in the middle of the experiment, and then declined.

Subject no. 11, which also responded to the bright light, commenced with a zero performance and, although the curve (see Fig. 12) was characterized by steep ascents and declines, the general trend was upward. *The steep ascent of 80 points after the 17t, day is noteworthy.* Perfection was reached on the 31st day after a period of no progress. *The three transpositions, including the test at the end* (see Table III) *failed to result in any increase in errors. The medium curve began at 100 percent, where it remained threh days.* It then descended rapidly and continued on the downward trend to the end of the experiment. The number of errors to the dim started at zero and increased rapidly to a peak on the fourth day, from which point it gradually decreased to the end of the experiment.

The curve of fish no. 12's performance (see Fig. 13), which also responded to the bright light, started at zero and maintained a fairly constant level up to the 36th day, meanwhile not going above 60 percent. The curve then mounted to a plateau at 80 percent—an irradiation pattern representing the maximum amount of insight that the animal then possessed into the problem. After four days the curve ascended to 100 percent and dropped only momentarily to 80 percent, when the first transposition was made. The next transposition, and the test at the end (see Table III), did not lower the curve from the 100 percent mark. Errors to the medium followed the bright responses very closely, but of course, decreased in number at the last when the bright curve went up. Errors to the dim increased rapidly at the beginning of the experiment and maintained a level that for the most part was higher than either the bright or the medium responses up to the 30th day, when they decreased rapidly to zero. The largest number of errors to the dim occurred toward the middle of the experiment, a fact that has been noted before.

Fish no. 13 was trained to respond to the medium light out of the three absolute intensities presented. The curve of its performance (see Fig. 14) is curious in that it started at 100 percent. It declined irregularly to a low point on the 16th day, then mounted very rapidly to 100 percent, evidence again for the rapid solution of the problem, when maturation is complete. The three transpositions, including the test at the end of the experiment, did not lower the curve from 100 percent (see Table III). A consideration of the errors shows that there were relatively few responses to the bright light and that these were confined to the first part of the experiment. Errors to the dim light began at zero and increased irregularly to a peak toward the middle of the experiment, from which point they decreased rapidly and remained few to the end.

The curve of fish no. 14 (see Fig. 15), which also responded to the medium light,

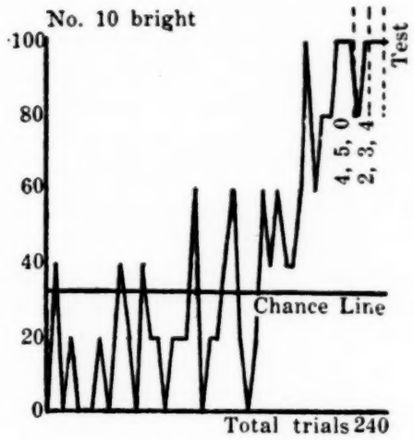


Fig. 11

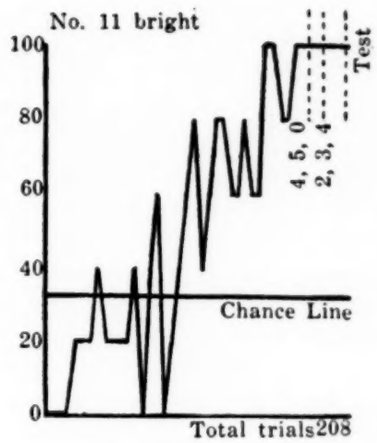


Fig. 12

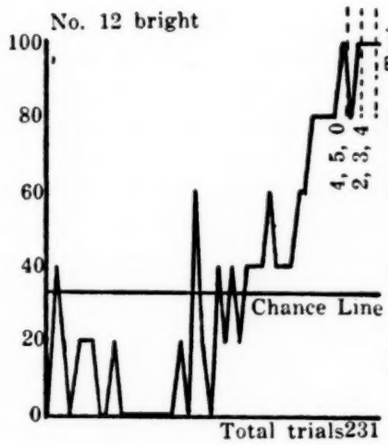


Fig. 13

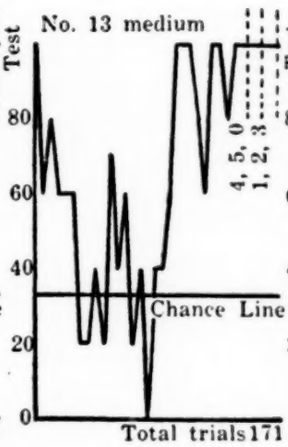


Fig. 14

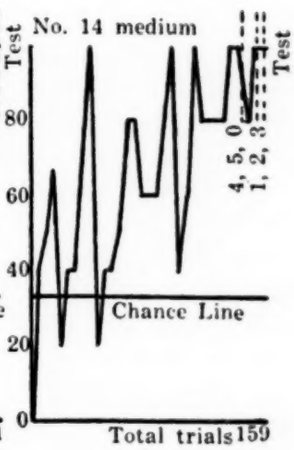


Fig. 15

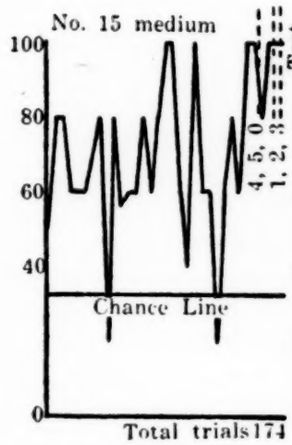


Fig. 16

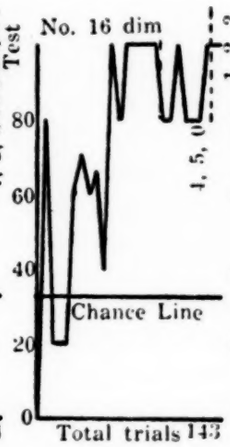


Fig. 17

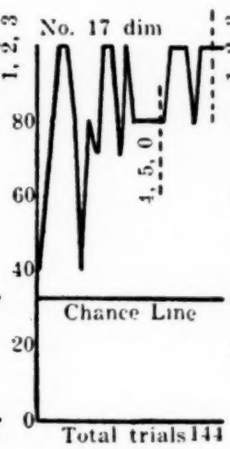


Fig. 18

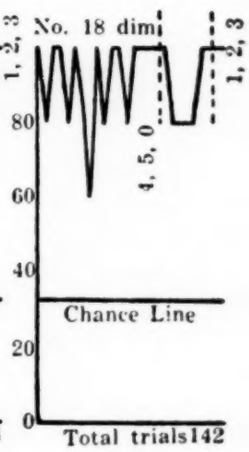


Fig. 19

commenced at zero, but mounted very rapidly. The curve was very irregular, ascending to a peak and then falling rapidly. It will be noted that the *four separate patterns of these ascents and declines were very similar.* The steep ascents again give evidence for the insight hypothesis and not for a blind trial and error behavior. The plateau, at 80 percent, just before perfection of the problem had been attained, is further evidence for the supposition that a brief or long period of maturation is often necessary before the animal can master a problem at first too difficult for it, and the fact that perfection was reached after this plateau had terminated seems to validate this supposition. The first transposition caused the animal to make one error, but the next transposition and the test at the end resulted in no errors (see Table III). The errors show that only two responses were made to the bright light in the entire experiment. Mistakes to the dim, therefore, run reciprocally to the learning curve.

Fish no. 15, which also responded to the medium light, showed a learning curve that was relatively high throughout. It was characterized by steep ascents and declines, *and again we see a characteristic pattern repeated in these fluctuations.* The general level of the curve is constant, until the decided drop on the 24th day. This is followed by a steep ascent to 100 percent (see Fig. 16). The first transposition caused a slight fall in the curve, the second, together with the test (see Table III), however, caused no increase in errors. The errors to the bright light maintained a fairly constant level throughout the experiment, except for the high number on the 24th day. Errors to the dim were slightly greater in number and rose to a peak on the ninth day and then gradually decreased.

Subject no. 16 was trained to respond to the dim light out of the three absolute intensities presented. This curve commenced at zero and mounted very rapidly through 80 points (see Fig. 17), then declined. Another steep ascent occurred on the tenth day, after which the curve never fell below 80 percent. The first transposition caused a slight decrease in accuracy, but the 100 percent mark was maintained after the second transposition. No test was given this sub-group at the end of the experiment. The errors show that there were only a few responses to the bright light and these were toward the first of the experiment. Mistakes to the medium light started at 100 percent and gradually decreased.

The curve of subject no. 17 (see Fig. 18), which also responded to the dim light, began at 40 percent and climbed rapidly to 100 percent. After the slight descent that followed, the record maintained a level for the most part of 80 percent or above to the end of the experiment. The plateau at 80 percent just before the transposition was made, probably indicated a 'growing stale' with respect to the problem, for immediately after the transposition, the curve mounted to 100 percent. The last transposition failed to lower the curve from the 100 percent level. The errors show that there was only one response to the bright light throughout the experiment. Mistakes to the medium light are, therefore, just the opposite, in percent, from the correct responses.

The curve for no. 18's performance (see Fig. 19), which also responded to the dim light, is very curious in that it started at 100 percent and fluctuated between that point and 80 percent (except for one descent to 60 percent) for the entire experiment. The first transposition did not cause an immediate decline, but one occurred the next day. The second transposition caused no decline from perfect performance. Errors to the medium light are very few, because of the high level of the dim curve. There was only one response to the bright light.

Control Group.—The results of the dynamic control group are given in Table IV. This table shows the choices made by

TABLE IV

RESPONSES OF THE CONTROL GROUP

Fish No.	Bright	Medium	Dim	Total
	Dynamic (1, 2, 3, 4, 5)			
1.....	0	8	12	20
2.....	5	3	14	22
3.....	5	12	3	20
4.....	2	9	9	20
5.....	3	10	7	20
Total..	15	42	45	102
	Static G (1, 3, 5)			
1.....	0	8	12	20
2.....	7	7	6	20
3.....	1	5	14	20
4.....	5	12	3	20
5.....	3	6	11	20
Total..	16	38	46	100
	Static H (2, 4, 0)			
1.....	1	7	12	20
2.....	1	5	14	20
3.....	2	7	11	20
4.....	3	8	9	20
5.....	3	8	9	20
Total..	10	35	55	100
	Static I (1, 2, 3)			
1.....	4	12	4	20
2.....	3	8	9	20
3.....	1	10	9	20
4.....	5	10	5	20
5.....	9	8	3	20
Total..	22	48	30	100
	Static J (2, 3, 4)			
1.....	5	9	6	20
2.....	5	13	2	20
3.....	2	12	6	20
4.....	2	14	4	20
5.....	1	10	9	20
Total..	15	58	27	100
	Static K (3, 4, 5)			
1.....	0	11	9	20
2.....	5	5	10	20
3.....	1	14	5	20
4.....	6	11	3	20
5.....	10	9	1	20
Total..	22	50	28	100

the various members of this group to the lights on the basis of the relative values: bright, dim and medium. It will be noted that in this group of control subjects, the greatest number of responses was made to the dim light, when the intensity as well as the position of the lights was changed after each response. Although the greatest number of responses was made to the dim light, the medium was responded to practically as many times. This would indicate a very slight preference for the dim light, but shows a marked avoidance of the bright light, no matter what its intensity.

Sub-group G of the static group responded to the light intensities 1, 3 and 5, merely changed in position. Table IV also shows the responses of this sub-group. It will be noted that the greatest number of responses were to the dim light, but that there was a greater difference between the responses to the dim and medium than in the preceding group. The bright light is again definitely avoided.

Table IV shows, again, that Sub-group H, which responded to the light intensities, 2, 4 and 0, made a striking majority of their responses to the dim light. The medium light was next in preference and the bright was chosen only a few times, ten out of 100 responses. The intervals between lights, as in the preceding sub-group, were one step, that is, a light intensity was skipped between each light used.

Sub-group I, the members of which were given the lights 1, 2 and 3, made the majority of their responses to the medium light (see Table IV). The dim light comes next in preference, but was favored only slightly to the bright light. In this case the light intensities used were not separated by steps, but were adjacent on our scale. However, these three light intensities were the three highest, on the scale used in the experiment.

The fish in Sub-group J, responded to the light intensities 2, 3 and 4, which were also adjacent, but fall in the middle of the intensity scale. In this case the majority of responses was over whelmingly in favor of the medium light, with the medium light in the center of the total range of intensities. The dim light was next in preference and the bright, last.

Sub-group K, was presented with the light intensities 3, 4 and 5, which were at the lower end of the intensity scale. The majority of responses was again made to the medium light, the dim coming next and the bright last. In this case, and in the sub-group preceding, 50 percent or more of the responses were made to the medium light.

Qualitative Results

Numerous observations are made in any learning experiment that cannot be treated quantitatively, nevertheless, they may be of greater importance, from the standpoint of comparative psychology, than the quantitative data. Many of the responses to be reported here occurred numerous times, not only in the same subject, but in all subjects. For this reason all but the most characteristic reactions are not described.

In none of the subjects was the response to a given light a sudden or direct process. The animal swam around the outer compartment before making a choice. In most cases it not only circled this compartment, but it actually nosed the various doorways, giving all indication that it was not responding to one light alone, but to that light in relation to the other lights in the stimulus-pattern. The animals frequently nosed the doorways that were next in intensity to the arbitrary standard prescribed by the experiment, as if 'comparing' the two lights before making a response. In many instances the fish, after a complete circuit of the choice compartment, darted suddenly through the correct doorway. This performance occurred in its most striking form when the animal was transposed to a new situation, such as the test situations presented.

In previous experiments of this kind on the goldfish,[5] when the animal had entered the wrong compartment and was punished, it would immediately dart into the correct compartment, before a change in the problem-situation could be made. Although this type of behavior was minimized by the apparatus used in this experiment (by the wire cage in particular) some of these responses still occurred. Statistically this corrected response was not included in the data of the present experiment. Only the first response of the two was counted. Nevertheless, from a qualitative standpoint this type of behavior is important since it is evidence of insight. Only in a very few instances did the animal swim from one wrong compartment to another wrong one and, what is more important, the second compartment was rarely directly across from the wrong one. This meant that the animal was obliged to *turn* in order to enter the correct compartment.

Another type of response frequently observed, toward the first part of the experiment, was movement toward and into a given doorway, where the fish backed in and out. This was especially true of the animals that responded to the bright light and may be interpreted as a 'fear' of that compartment. That is, the bright light set up a high tension within the animal, one that offset the tension toward the goal. As a consequence the animal did nothing, temporarily.

On the ninth day fish no. 1 (bright) made a devious circuit of the choice-compartment, nosed the dim, bright, medium, and then chose the dim light (wrong, see Fig. 20, no. 1). On the 11th day this animal swam directly toward the dim light, nosed it, backed away, nosed it again and then darted to the bright (correct, see Fig. 20, no. 2).

---

[5] Perkins, F. T., and Wheeler, R. H., *Loc. cit.*

When the combination 4, 4, 4 was given in the test on the 51st day, this animal circled around in the center of the choice-compartment, but would make no response.

Fish no. 2, on the 17th day, made a circle of the choice-compartment, nosed the dim, bright, and then chose the medium (wrong, see Fig. 20, no. 3).

On the third day fish no. 3 nosed the medium light several times, nosed the dim, swam a complete circuit of the choice-compartment and then chose the bright light (correct). On the fifth day this animal made two and one half circuits of the choice-compartment, nosed the bright and dim several times and then darted to the dim compartment (wrong, see Fig. 20, no. 4). This fish made a zigzag circle of the choice-compartment, on the ninth day, nosed the dim several times, then went to the medium (wrong, see Fig. 20, no. 5). On the 11th day it headed toward the medium light, circled around and chose the bright (correct, see Fig. 20, no. 6). No. 3 headed for the dim light on the 12th day, nosed the bright, chose the medium (wrong, see Fig. 20, no. 7). Two days later this subject headed for the bright light, nosed it, swam to the medium, swam across the choice-compartment to the dim, turned and chose the bright light (correct, see Fig. 20, no. 8). On the 29th day it swam toward the dim compartment, turned and swam to the center of the choice-compartment, headed for the bright, backed away, then darted into the bright compartment (correct, see Fig. 20, no. 9). On the 30th day this animal circled around the choice-compartment an astonishing number of times, nosing all the compartments, finally choosing the bright light (correct, see Fig. 20, no. 10). On the test on the last day this animal's behavior was markedly different; it was less sure—more nosing; more circling—slower in responding. When the combination 1, 2, 1 was presented, it was very slow in choosing, nosed both no. 1s several times and finally chose the 1 on the right side.

Fish no. 4, which was given the medium light darted for the dim light on the 12th day, backed away, darted for the bright light, backed away and chose the medium (correct, see Fig. 20, no. 11).

On the 14th day, no. 5 headed for the bright light, circled back to the middle of the choice-compartment, spiraled to the dim light, passed on, and chose the medium light (correct, see Fig. 20, no. 12).

Fish no. 6 headed straight for the dim light on the ninth day, nosed it, swam back toward the center of the choice-compartment, and then darted into the medium doorway (correct, see Fig. 20, no. 13). On the 20th day this subject swam to the dim light, then to the bright light, nosed it, backed away and chose the medium (correct, see Fig. 20, no. 14).

On the test, fish no. 7 would not choose when the 1, 1, 1 combination was presented, but darted in zigzag fashion all around the choice-compartment. When the 5, 5, 5 combination was presented, this subject made a complete circuit of the choice-compartment, starting at the left side and finally chose the left light (see Fig. 20, no. 15). It did not respond on the 3, 3, 3 combination, but headed to the left light, nosed it, backed away, nosed the middle, backed away, nosed the right hand one, backed away, swam to the left again and then back to the middle of the choice-compartment (see Fig. 20, no. 16) and stopped.

Subject no. 9 headed for the medium light on the eighth day, turned and circled back, swam to the bright light, nosed it, returned to the center of the choice-compartment and then chose the dim (correct, see Fig. 20, no. 17). On the 22nd day this animal swam toward the bright compartment, nosed it, turned to the medium light, nosed it, then swam straight across the choice-compartment and selected the dim (correct, see Fig. 20, no. 18). On the test when the 4, 4, 4 combination was presented, this fish swam around the choice-compartment very rapidly, but did not choose.

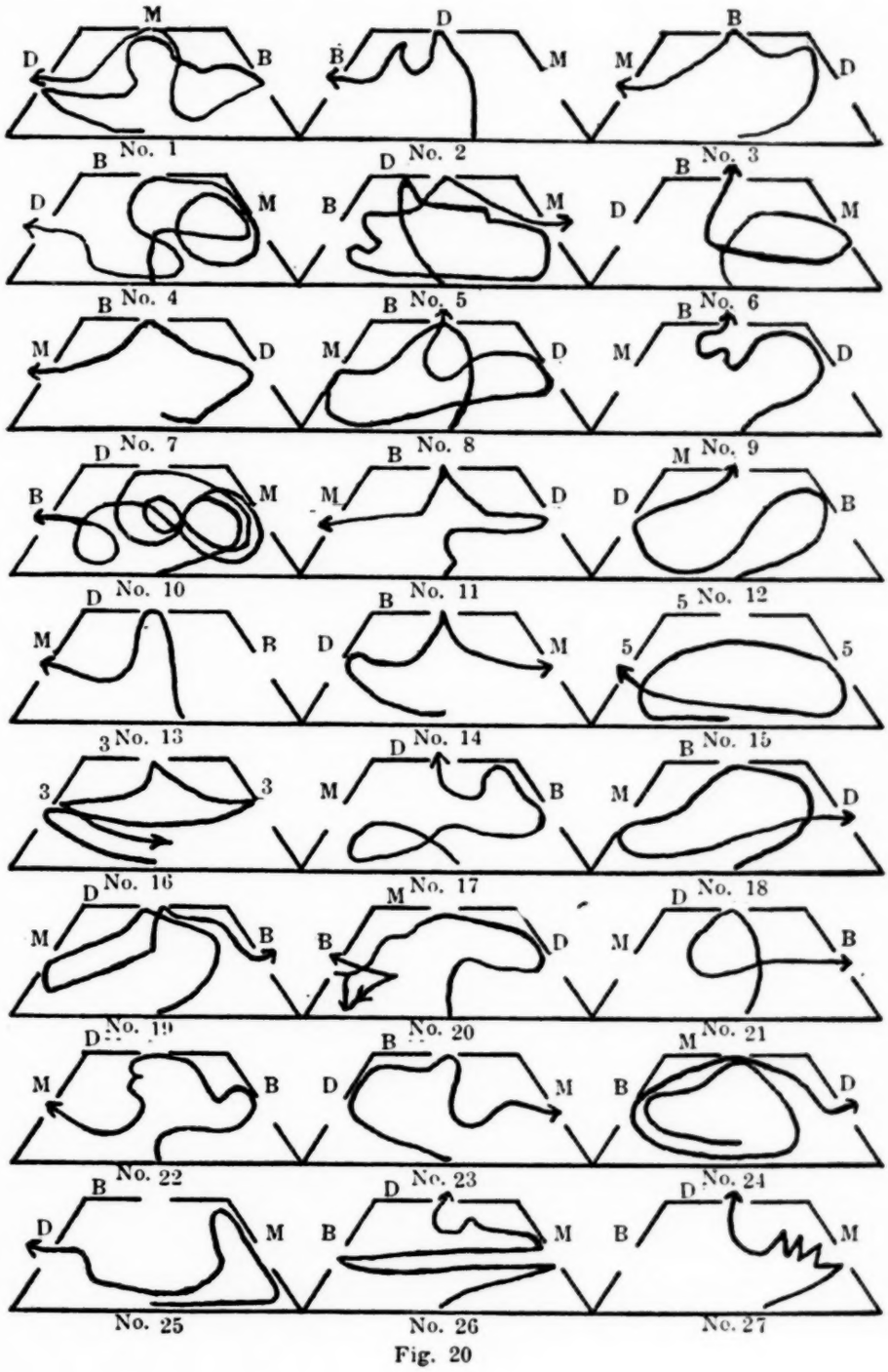

No. 1
No. 2
No. 3
No. 4
No. 5
No. 6
No. 7
No. 8
No. 9
No. 10
No. 11
No. 12
No. 13
No. 14
No. 15
No. 16
No. 17
No. 18
No. 19
No. 20
No. 21
No. 22
No. 23
No. 24
No. 25
No. 26
No. 27

Fig. 20

On the 20th day no. 10 swam as if to enter the bright, but turned and nosed the dim, nosed the medium, nosed the dim again, then chose the bright (correct, see Fig. 20, no. 19). On the 38th day this subject swam to the dim light, nosed it, nosed the medium, swam to the bright compartment, nosed it, backed away, turned, backed again, turned again and chose the bright (correct, see Fig. 20, no. 20). On the test this animal was markedly disturbed when the 4, 4, 4 combination was presented and swam in circles, never choosing.

On the 17th day no. 11 swam to the dim compartment, nosed it, swam to the medium light, nosed it, swam to the bright, nosed it, then made a complete circuit of the choice-compartment and entered the bright doorway (correct, see Fig. 20, no. 21).

On the 11th day no. 13 headed for the bright light, turned and nosed the dim light, turned, swam toward the center of the choice-compartment, vacillated forward and back in front of the medium light, then chose it (correct, see Fig. 20, no. 22). On the test this subject turned its head away from the lights when the 1, 1, 1 combination was presented and remained in this position, not making any choice.

Fish no. 15, swam to the dim compartment on the 16th day, turned at the door, nosed the bright compartment, swam to the center of the choice-compartment and then darted to the medium light (correct, see Fig. 20, no. 23).

On the 11th day, fish no. 16 swam to the bright compartment, nosed it, turned toward the medium light, then made a circuit of the choice-compartment, nosed bright and medium again, then chose the dim (correct, see Fig. 20, no. 24).

On the 11th day no. 17 swam along the outer wall of the choice-compartment to the medium light, swam on to the bright light, then turned around sharply and went back to the front wall and turned to the dim light (correct, see Fig. 20, no. 25).

On the 12th day no. 18 swam straight to the medium light, backed away and darted to the bright light across the choice-compartment, darted back to the medium and then chose the dim (correct, see Fig. 20, no. 26). On the same day it headed for the medium light, backed away a small distance, then darted forward, backed away, and darted forward again. It did this four times in front of the medium light, each time, however, approaching nearer the dim light. It then darted through the dim doorway (correct, see Fig. 20, no. 27).

In all of the subjects another mode of response revealed itself as a result of the method used. The animals learned to come out of the correct compartment at the end of about ten seconds and to swim into the raised cage, which had been placed for the next trial. The animals also spent less time in what might be called exploratory movement as the experiment progressed, spending all their time in the choice-compartment responding directly to the lights.

The material presented in this section indicates very strongly that the fish were not reacting to isolated stimuli, but that they were responding to a total stimulus-pattern. Even more important than this is the fact that the animal was *making* a *choice;* the fish itself was the important factor in the response. The animal was not merely present in the experimental situation, responding passively, in a tropistic fashion. *It was responding to a certain light, because of the value of that light with respect to the total stimulus-pattern.*

Discussion

Comparison of Groups.—From the graphs of the experimental group the main generalization is evident that there was a marked avoidance of the bright light; the bright light-

problem seemed to be the most difficult. The data from the control group would corroborate this fact. If we tabulate the total number of trials necessary for the mastery of the various problems, we then have a basis by means of which to compare the various groups (see Table V). In both the

TABLE V

NUMBER OF TRIALS FOR EXPERIMENTAL GROUP

Dynamic Group

Fish No.	No. of Trials	Average Trials	Total
Bright			
1	257		
2	241	219.3	
3	160		658
Medium			
4	169		
5	183	180	
6	188		540
Dim			
7	133		
8	106	133.3	
9	161		400

Static Group

Bright			
10	240		
11	208	226.3	
12	231		679
Medium			
13	171		
14	159	135	
15	174		504
Dim			
16	143		
17	144	143	
18	142		429
			3,210

static and dynamic groups the bright subjects required the greatest number of trials to learn the problem. The fish responding to the medium light took the next largest number of trials and the dim the least. In each group the ratio of 6/5/4 held for the three stimuli, bright, medium and dim, respectively. In every case, however, except the medium, the number of trials required for the static group exceeded the number for the dynamic group. This fact indicates that

a changing configuration brings out or exemplifies the problem to be learned. In a changing pattern the relationships are more easily seen.

Those fish that responded to the dim light seemed to learn the most rapidly. It would, therefore, look as if the dim light was preferred. In previous observations it had been noted that the goldfish seemed to prefer the medium light,[6] no matter what its absolute intensity. In order to find some clue as to this discrepancy, the responses of the control group were examined in more detail. The reactions of each fish were divided into four groups of five trials, and the responses to the three lights recorded separately (see Table VI).

TABLE VI

AVERAGE OF GROUP PERCENTAGES OF JUDGMENTS TO EACH LIGHT

	Dynamic (1, 2, 3, 4, 5)			Static (1, 2, 3) (2, 3, 4) (3, 4, 5)			Static (1, 3, 5) (2, 4, 0)			Total Average		
	B	M	D	B	M	D	B	M	D	B	M	D
1.....	24	44	32	25	51	24	14	46	40	31	47	32
2.....	24	32	43	22.5	51	26.5	12	26	62	19.5	36.3	43.8
3.....	4	44	52	15	49	36	14	32	54	11	41.8	47.3
4.....	4	48	48	16	57	27	12	38	50	10.8	47.8	41.8
Av.....	15	42	45	19.66	52.66	28.3	13	36.5	50.5	18.07	43.22	41.2

Legend: B, bright.
M, medium.
D, dim.
Av., average.

Table VI shows that the free responses to the bright light diminish with time, thus indicating why the fish that responded to the bright light in the experimental group were slower to learn. The free responses to the dim light *increase* with time, and correlate with the faster learning noted in the animals that chose the dim light in the experimental group. The same quantitative data yield further evidence for the supposition [7] that responses to the lower end of the intensity scale represent less difficult choices than those to the upper end

[6] Perkins, F. T. and Wheeler, R. H., *Loc. cit.*, 40.
[7] Perkins, F. T. and Wheeler, R. H., *Loc. cit.*, 39.

of the intensity scale. It will also be noted that, *on the whole, the control group preferred the medium light.* This fact did not seem to harmonize with the results of the experimental group. A further analysis of the reactions of the control sub-groups was made (see Table VI). It turned out that the control group responding to the changing light pattern, with the five intensities, did not seem to prefer either the medium or the dim light, although a few more responses were made to the dim. In the static group, the wide-interval sub-groups, or those responding to lights separated by one interval, made the majority of their responses to the dim light, in keeping with the experimental group. The control sub-groups that responded to three lights, adjacent in intensity, *made a large majority of their responses to the medium light. This means that when the stimulus lights are separated by a sufficient interval of intensity, the animal chooses the dim light, but when the lights are nearer together in intensity the animal prefers the medium light.*

Average curves are here included for each sub-group of the experimental group. Curves for the fish choosing the bright light (see Figs. 21 and 22) in both the static and the dynamic groups, mount slowly throughout the experiment. These curves are very irregular, but are somewhat similar in nature. The curve for Sub-group A (Fig. 21) reached a high point in the middle of the experiment, caused by the one subject that learned the problem much more rapidly than the other two.

The curves for the two medium sub-groups (see Figs. 23 and 24) of the static and dynamic, commence with a much higher percentage of accuracy than the bright, and then decline to a level that is maintained to the middle of the experiment (about 18 days). From this time on the curves mount very rapidly, but are very irregular. The curves are much shorter than those for the 'bright' sub-groups.

The curves for the 'dim' sub-groups (see Figs. 25 and 26), both start at about the same level as the 'medium' fish, but mount from this point instead of declining. They both remain fairly high throughout the experiment, but are quite irregular. Both curves are very short.

Interpretation of Results.—It has already been mentioned that the experimental group, as well as the control group, seemed to avoid the bright light. The experimental animals, trained to choose the bright light, gave the longest learning curves and the other sub-groups made very few errors to the bright light. Also, the members of the control group made

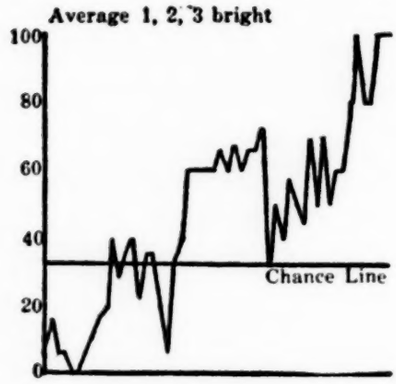


Fig. 21

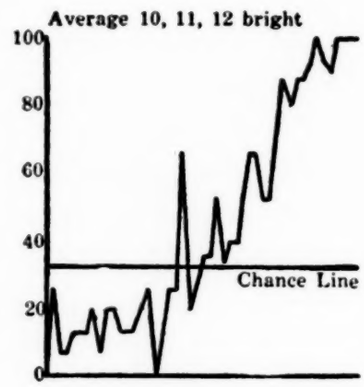


Fig. 22

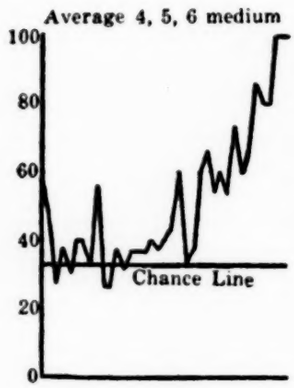


Fig. 23

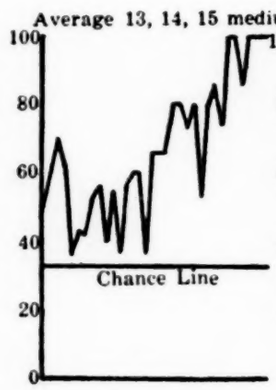


Fig. 24

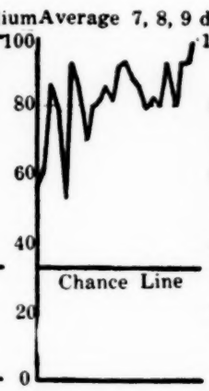


Fig. 25

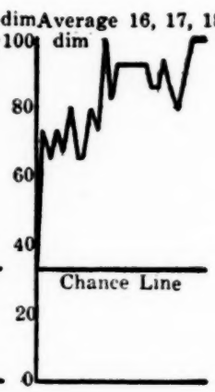


Fig. 26

very few free choices to the bright light. The bright light in any pattern of three intensities, then, must represent in these animals a certain high tension, that is, it sets up a high stress, which must be resolved in line of some member of the pattern that represents a low stress. Here we are making use of a general law, the law of least action, which applies to all behavior [8] and states that in any energy system where there is a difference of potential, activity will take place to

[8] See Wheeler, R. H., The Science of Psychology, Crowell, 1929, for a discussion of the applicability of this law to behavior.

resolve the high potential over the most direct route in time in the direction of the low stress or potential. In this case the bright light represents the high stress at the outset of the experiment, whereas the lights toward the lower end of the intensity scale represent the low stress toward which activity will take place. However, in previous experiments the light of medium intensity was preferred,[9] which would indicate that the dim light also represented a high stress in the behavior pattern. In this experiment this preference of the medium light did not appear to hold when the learning curves were examined, but a critical examination of *all the control groups shows that on the average there is a preference for the medium light at the outset of the experiment.* The medium light does, then, at first represent the low stress in the configuration, toward which the high stresses set up by the bright and the dim lights will be resolved. As the experiment proceeded, other goals than preference were set up which conditioned the learning of the problem. In the case of the fish that were to respond to the bright light, the stress toward avoidance set up by the bright light was much greater than that set up by the dim so that the food-goal in case of the former had a greater stress to overcome before the animal could learn the problem. This would also mean that as soon as the fish had eaten, the stress toward the food-goal would diminish in intensity and the animal would make more errors, a fact testified to by the long irregular learning curves in case of the bright.

We have noted that the learning curves seem to point to the fact that the dim light was preferred to the other two in both experimental groups, but that the results of the control group gave evidence of no such preference at the outset. It should then be expected that the 'medium' errors and learning curves in the experimental group would start relatively high. The curves of the entire experimental group show that ten out of the eighteen fish commenced choosing the medium light above a chance percentage. That the other eight subjects did not choose the medium is probably due to the fact that training at once destroyed the mere preference-

[9] Perkins, F. T. and Wheeler, R. H., *Loc. cit.*, 40.

reactions. The control group's results indicate that the medium is preferred in the first ten trials and that the dim is much less attractive at first, but the preference for the dim light *increases* as time goes on, so that at the end of 20 trials it is preferred to the medium light. The reason for this is probably that other goals, yet undetermined, are set up at this time, sooner than in the original experiments,[10] because of better controlled conditions. The fact that the free responses to the medium light diminish in number as time goes on accounts for the drops in the learning curves in the first part of the experiment and is corroborated by results from the original experiments, where it was found that the medium curves were high at the outset, but became lower as the experiment continued.[11]

Another factor to be considered in this connection is that the dim light seemed to be preferred in the control group when the lights were separated by a step in intensity. However, examination of the free choices divided into groups of fives (see Table VI), shows that *even under these conditions the medium is preferred at the outset*, whereas the dim is preferred later on. *In the case of the sub-groups that responded to a stimulus-pattern in which the lights were not separated by steps, the medium light is preferred throughout.* This means that when the lights are separated by steps the other goals, probably food, are set up more quickly because the configurational response is more easily formed, that is, the relationships are more obvious; whereas, if the lights are close together, the responses to the other goal are delayed in the course of forming more difficult configurations.

A preference for the medium light at the first of the experiment indicates not only a response in the line of least action, which has already been mentioned, but it also means that the animal has made a balanced response, of the simplest possible type, to a stimulus-pattern that involves the entire range of light intensities. That is, the response to the medium light might be likened to a response in the human subject to a

[10] Perkins, F. T. and Wheeler, R. H., *Loc. cit.*

[11] Perkins, F. T. and Wheeler, R. H., *Loc. cit.*, 41.

symmetrical arrangement of stimuli. The bright and the dim light represent the two extremes of the pattern, while the medium light represents the mean or center. Responding to the mean in relation to the extremes, makes of the neuromuscular configuration a symmetrical pattern. That the symmetrical pattern predominates is again evidence of the law of least action, for the symmetrical pattern represents the simplest and most stable pattern and, therefore, the one that requires the least energy to form and maintain. This interpretation would also provide the reason why the responses to the medium light diminish as time goes on. The setting up of the food-goal, as a member of this particular configuration, would upset the balance of potentials in the pattern and therefore the response to the medium light would no longer be a symmetrical configuration.

The concepts of insight and maturation have been mentioned from time to time in the discussion of the behavior of the individual fish. By insight is meant "(*a*) transposition of the general properties from one situation to another, (*b*) response to a complex, total situation the details of which have been altered without changing their relationships in the total structure, (*c*) response to parts in the light of the whole, (*d*) configurational or structured response, (*e*) reacting to a complex situation in a way that might be called adequate or 'sensible'." [12] In the learning curves we have seen evidence for this type of response in the frequent steep ascents. Insight, as we shall use the term, refers to the problem that we have set up for the animal, although the animal's entire behavior, if we consider each response separately, is insightful in character or the animal would not respond at all. The insight into any particular problem is the result of a maturation process,[13] a general growth process, which is conditioned by stimulation.

[12] Perkins, F. T. and Wheeler, R. H., *Loc. cit.*, 1. Insight is used here in the sense that Köhler used it. Köhler, Wolfgang, *The Mentality of Apes* (E. Winter, Tr.) London: Kegan, Paul, 1924. Helson, H., Insight in the White Rat. *Jour. Exper. Psych.*, Vol. 10, 1927, has also applied the concept to animal behavior.

[13] See Wheeler, R. H., The Science of Psychology, for a discussion of the concept of maturation.

We have noted in our examination of the learning curves that the curves ascend in volleys, or follow a certain pattern that is characteristic of the animal. This pattern in the learning curve may represent a picture of the maturation process that is going on and may give us evidence as to how this can be controlled for effective learning. These patterns are usually preceded by a low point in the curve, or a plateau, which Snoddy has called an irradiation pattern.[14] This plateau represents in animal learning an inability to see a detail in relation to the total situation, an insufficient insight into the problem presented. The patterns or volleys in the learning curves show us how maturation proceeds, from a low point or a plateau, in a similar fashion through repeated increments of development. We may then envisage the learning process as going on in certain multiples of a maturation pattern or rhythm that is characteristic for the animal.

As has already been noted, among the experimental animals, the dynamic group was superior to the static group. That this is true seems to substantiate a general principle in psychology, namely, that the animal responds to a total stimulus-pattern and if to some detail of that pattern, to that detail in its relation to the other details of the pattern, or, in other words, to the pattern taken as a whole. This principle is known as the law of configuration,[15] and meant, in the present experiment, that the animal was responding to the pattern of three lights and if to the bright light, to that light in its relation to the medium and the dim. In fact the quality of 'brightness' possessed by the bright light is derived from the pattern of which it is a member. The animal that responded to the bright light would then see the relationship of bright to the other two all the more plainly if the absolute intensities of the individual lights were constantly changing, because this method emphasized the relationship and submerged the definiteness of the absolute intensities. Then we would expect just what actually did occur, that the members of

[14] Snoddy, G. S., An Experimental Analysis of a Case of Trial and Error Learning in the Human Subject, Psych. Mono., Vol. 20, No. 124.

[15] See Wheeler, *Loc. cit.*, for a general application of the law of configuration.

the dynamic group would learn the problem more rapidly than the static group. Another aspect of this response to the total situation, so beautifully demonstrated by the dynamic group, is the fact that they were learning the task by the whole method, that is, they learned all of the problem at once, whereas the other group were forced to learn it by parts (serially). This would mean, then, that the whole-method group could see the total pattern throughout the experiment and thus could learn more rapidly.

The qualitative data from the various groups illustrates still more conclusively that the animals were responding to a total pattern. The animals would not swim directly to a given light, but would circle the choice-compartment and literally 'look over' the situation before making a response, thus indicating that no one member of the stimulus-pattern had unique properties as a goal in itself, but derived those properties from the rest of the pattern.

Summary

1. The goldfish is capable of learning to choose one of three absolute light intensities and to transpose both up and down the intensity scale, *without added training.* This animal is also able to choose one light on the basis of its relationship to two others, when the intensities as well as the positions are changed after each trial.

2. A constantly changing pattern of light intensities is learned more quickly than a stereotyped pattern.

3. Under all the conditions of this experiment the goldfish prefers the medium light at the outset, but if the lights are separated relatively widely in intensity a preference develops for the dim light.

4. The results of this experiment demand an interpretation of the behavior of the goldfish in terms of insight, rather than in terms of trial and error.

5. This animal responded to total situations in accordance with the law of configuration.

6. The animal's response to the medium light at the outset represents behavior in the line of least action. The response

is a simple, balanced pattern structured with respect to a total range of intensities of stimulation, of which the average or medium intensity represents the lowest stress.

7. The general notion of maturation is applied in explanation of the development of insight into the experimental problem. The characteristic patterns or volleys through which the learning curves ascend may be objective pictures of the way maturation takes place in the individual animals.

8. The results of this experiment confirm and supplement a previous study of behavior in the goldfish by Perkins and Wheeler.

(Manuscript received September 15, 1930)

THE MEMORY VALUE OF SIZE *

BY THEODORE KARWOSKI

Dartmouth College.

The following experiments were undertaken in order to determine the memory value of size under conditions in which size alone is the variable. In the average situations involving the factor size, as in advertising, for example, other factors such as attention, feeling-tone, established associations etc., are inseparately linked with magnitude.

In these experiments two digit numbers were selected but the easily remembered combinations, were omitted, as for example, 10–20, 25, 50, 75, 22, 34, 43 etc. Care was taken, also, to avoid placing readily associated numbers near each other in the series. In this way six series of 12 two digit numbers were composed. These six series were triplicated, so that one set of six was printed in small, another in medium and a third in large size. The heights of the three magnitudes were related as 4 : 6.5 : 9.5 mm. The ratios of small heights to medium and of medium to large were 1.6 and 1.46 respectively.

The observers were seated comfortably, at such a distance from the slot in the mechanically operated memory drum, so that they could see clearly all the material exposed at the slot. This distance was suited to the individual subject. The numbers appeared, one by one, in the slot at the rate of one per second.

The subjects were given instructions to keep their eyes continually fixed at the slot during the experiment. Immediately after presentation of the series they reported what numbers they saw. The visual and passive attitude was required; the subjects were asked to avoid forming associations between numbers as well as to refrain from recitation and

* The author is indebted to Professor F. H. Allport for his interest and encouragement in the early stages of this study.

articulation during the exposure of numbers. This attitude was required in order to limit the research as much as possible to size in terms of vision. The passive attitude retarded the learning process to such an extent that it was possible to use these series of numbers repeatedly without overlearning.

Experiment on Memory Value of Absolute Size

In this experiment eleven subjects, attending courses in psychology at the Harvard summer school, were used. They were presented the three sets of six series of numbers, in such a manner, that practise effects were equalized as between the three different sizes. After preliminary practise tests, each subject sat for three sessions. During each period six series were presented, of which two were all small, two were all medium, and two were all large sized numbers. The results are represented in the following table.

TABLE I

Number of Two Digit Numbers Recalled

Observer	Large	Medium	Small
1	26	24	27
2	24	25	26
3	19	25	23
4	18	18	20
5	20	20	25
6	18	17	17
7	9	11	8
8	23	13	14
9	18	18	13
10	29	30	33
11	20	20	17
	224	221	223
Av.	20.37	20.09	20.27
A.D.	3.73	4.21	5.84

The above results show that absolute size has no memory value. Newhall and Heim [1] investigated the memory value for *absolute size* in advertising and have obtained similar

[1] Newhall, S. M., and Heim, M. H., Memory Value of Absolute Size in Magazine Advertising. *Jour. of App. Psychol.*, 1929, XIII, pp. 62–75. In this article the writers define *absolute* size as extent of stimulation in the total field; the amount of difference or size contrast in stimulation is peculiar to *relative* size.

results with magnitudes of quarter, half, and full page of the average magazine.

Experiments on the Memory Value of Relative Size

This experiment was conducted under conditions similar to those of the previous experiment. The same numbers were used but instead of triplicate sets which were either all large, all medium, or all small numbers, triplicate sets were made such that the three sizes were distributed throughout each series of numbers. Thus in a single series of six which made up a set, the three sizes were represented equally—four large, four medium, and four small. These triplicate sets were so constructed that numbers that had been large in one set, were made small in the second and medium sized in the third set. Hence each number appeared in the same position in the series in all three sizes. The distribution of sizes within the series was determined by chance. Six subjects took the tests; four graduate students in psychology and two senior students. These observers were presented one complete set per experimental hour. Two of the six series constituting a set, were presented once, two twice, and two four times. Recall was required immediately after presentation. The series and sets were so spaced in time as to insure equalization of practise effects.

The following tables show the results in terms of all the numbers recalled, the percent contributed by large numbers (L), by medium (M), and by small (S). The columns labelled L–M, L–S, and M–S show the difference in terms of percent of the sizes compared. Plus sign indicates that the larger of the two sizes compared is the better recalled, minus indicates that the smaller is better remembered. The actual number of times large, medium, and small sized numbers were recalled is shown after the label "Number Recalled". Since each table represents a total of 1296 numbers presented, of which 1/3 were large, 1/3 medium, and 1/3 small, some idea of the degree of learning may be obtained. The memory value of size was determined by arbitrarily allowing the value for large size to be 1.00. Table II refers to the results of one

presentation of each series, Table III, two presentations, and Table IV, four presentations.

TABLE II

ONE PRESENTATION

Observer	L	M	S	L–M	L–S	M–S
Hul	49.39	23.18	27.53	+26.21	+21.86	− 4.35
Find	44.61	27.69	27.69	+16.92	+16.92	.00
Mark	40.54	27.02	32.43	+13.52	+ 8.11	− 5.41
Fraz	42.87	34.06	23.07	+ 8.81	+19.80	+10.99
All	41.57	26.96	31.46	+14.61	+10.11	− 4.50
Gold	39.25	34.57	26.16	+ 4.68	+11.50	+ 8.41
Aver	43.04	28.91	28.05			
No. recalled	211	146	138			
Mem. value	1.00	.692	.654			

TABLE III

TWO PRESENTATIONS

Observer	L	M	S	L–M	L–S	M–S
Hul	54.54	19.48	25.97	+35.06	+28.57	− 6.49
Find	40.23	28.73	31.03	+11.50	+ 9.20	− 2.30
Mark	40.21	30.43	29.34	+ 9.78	+10.87	+ 1.09
Fraz	32.98	40.20	26.80	− 7.22	+ 6.18	+13.40
All	40.40	24.24	35.35	+16.16	+ 5.05	−11.11
Gold	34.92	31.74	33.33	+ 3.18	+ 1.61	− 1.59
Aver	40.54	29.13	30.30			
No. recalled	230	171	177			
Mem value	1.00	.743	.769			

TABLE IV

FOUR PRESENTATIONS

Observer	L	M	S	L–M	L–S	M–S
Hul	42.72	29.09	28.18	+13.63	+14.54	+ .91
Find	35.13	35.13	29.73	.00	+ 5.40	+ 5.40
Mark	37.11	32.98	29.89	+ 4.13	+ 7.22	+ 4.09
Fraz	34.61	30.00	35.38	+ 4.61	− .77	− 5.38
All	38.18	34.54	27.27	+ 3.64	+10.91	+ 7.27
Gold	35.38	35.38	29.23	.00	+ 4.15	+ 4.15
Aver	37.17	32.52	29.94			
No. recalled	255	226	207			
Mem value	1.00	.886	.782			

The above data lead to the conclusion that a memory value for relative size has been obtained as between the largest and either of the two smaller sizes. Memory value for the larger of the two smaller sizes was not obtained even though the ratio between these sizes is about the same as the ratio between the two larger magnitudes.

The tables also show that the repetition of the series diminishes the memory value of size. With four repetitions the condition of equality with respect to memory value is approached.

The validity of these results may be ascertained by inspection of the data. If all observers show the same general tendency the probability that the tendency is significant is great. In the tables the columns labelled L–M, L–S, and M–S indicate by their sign the tendency of each subject. It is clear that all the subjects recall the largest size better than either the medium or the small size. There is one exception in Table III, and two in Table IV. The tendency, however, is quite marked. The column M–S shows a diversity of tendencies. Some remember the medium better than the small size, others vice-versa, and still others fail to show any difference. That this is not a matter of individual preference for size is denoted by the fact that the same individuals do not always show the same sign in all the tables. Hence it is likely that these differences between individuals are due to factors other than size difference.

The analysis of the above data may be refined statistically by studying the unreliabilities of the averages for large, medium and small magnitudes. Since the passive manner of memorizing, as used in these experiments, gives a relatively level practise curve the conditions are sufficiently stable to lump the scores of the six observers and thus obtain frequency curves in terms of the number of two digit numbers recalled. In each series there were four large, four medium, and four small sized numbers, and so the number of times any magnitude was recalled per memory exercise varied from 0 to 4. There were 108 such exercises for experiments of 1, 2, and 4 presentations of a series. The following table gives the

frequency measures, the averages, the difference between the averages for large and medium and small magnitude, the P.E. of the difference and the ratio of the difference to the P.E. of the difference. The tables show that the differences between the averages of the large and medium sizes are statistically reliable because the ratio of the difference between the averages to the P.E. difference is at least 3 : 1. The difference between the averages for the medium and the small sizes are not statistically reliable; apparently these magnitudes are responded to indifferently. With four presentations of the memory series the unreliability of the averages is greater than with one and two presentations and this apparently means repetition tends to obliterate the memory value for size.

TABLE V

THE UNRELIABILITY OF THE DIFFERENCES FOR RECALL OF MAGNITUDES

One Presentation

	No. recalled					Average	Difference	P.E. Difference	Difference / P.E. Difference
	0	1	2	3	4				
L....	10	25	40	30	3	1.916			
M....	25	36	37	11	1	1.361	.555	.0905	6.130
S.....	24	46	28	8	2	1.240	.121	.0922	1.312

Two Presentations

	No. recalled					Average	Difference	P.E. Difference	Difference / P.E. Difference
	0	1	2	3	4				
L....	4	20	51	24	9	2.128			
M....	19	30	39	17	3	1.585	.543	.0927	5.877
S....	16	33	38	17	4	1.629	.046	.0972	.473

Four Presentations

	No. recalled					Average	Difference	P.E. Difference	Difference / P.E. Difference
	0	1	2	3	4				
L....	3	20	36	36	13	2.333			
M....	7	25	34	35	7	2.092	.241	.0932	2.930
S.....	13	22	40	26	7	1.982	.110	.0973	1.133

Experiment on the Span of Cognition as Effected by Magnitude

This experiment was undertaken in order to discover what other factors are involved in the above conclusion, that the memory value for size does not correlate with the order of magnitudes. Although the observers reported that all the memory material was clearly visible and easily cognized, yet differences in perceptibility between sizes may not be apparent though present.

The method of 'constant stimuli' was adopted as a mode of procedure. Cards containing numbers ranging from 1 to 8 digits were exposed for 1/12 of a second; the subject reported what he saw in the manner of the ordinary 'span of attention' experiment. The cards were triplicated. One set contained stimuli of large size, another of medium size, and a third of small size. The numbers were the same as those used in the previous experiments.

The figures were arranged on the cards into the form of a semi-circle. In the center of this circle was a pinhole, which was used as a fixation point. These cards were placed at the farther end of a large wooden box. The subject looked into this darkened chamber from a distance of 90 cm. from the observers head rest. Under these conditions the subject could see nothing except the beam of light penetrating the pinhole and serving as a fixation point. The subjects eyes were outside of the box, therefore light adapted.

At a given signal the stimulus card was flooded with light from a 100 watt Tungsten Lamp contained in an adjacent box, which completely shut out the light except when the four inch Ilex shutter was released. In this manner tachistoscopic conditions were obtained with definite fixation and foveal stimulation. Thus the retinal stimulation in this experiment was similar to that of the memory experiments.

Four new subjects were available for this experiment. They all reported that the numbers could be easily seen and cognized at the distance of 90 cm. To test these statements the distance at which the small numbers were just visible was determined for each subject. For Ds. this distance was

337.82 cm; Br., 312.42; Bu. 297.18; and for We., 190.5 cm. All subjects except We. had normal vision; the latter was myopic. The results of 20 presentations of each stimulus card in each size are embodied in the following table and graphically represented in Fig. I. The numbers in the tables stand for the percent of correct responses.[2]

TABLE VI

(SMALL SIZE)

Observer	1 digit	2	3	4	5	6	Limen
Br.	100	74	57	43	35	4	3.67
Ds.	100	84	71.4	56.5	33.3	4	4.19
Bu.	90	55.5	8				2.
We.	52	32	8				1.

TABLE VII

(MEDIUM SIZE)

Observer	2	3	4	5	6	7	Limen
Br.		100	87.5	60	20	5.5	5.280
Ds.	100	96	83.3	64	32	6	5.402
Bu.	96	56	22	0			3.243
We.	92	56	48	22	0		3.730

TABLE VIII

(LARGE SIZE)

Observer	3	4	5	6	7	8	Limen
Br.		96	83	42	16.5	0	5.833
Ds.	96	92	84	32	6		5.601
Bu.	96	84	22	5	0		4.545
We.	96	92	44	12	0		4.833

The curves in Fig. I show that the range of cognition[3] varies with the size of the stimuli. The difference is greater between the two smaller magnitudes than between the two larger ones. These curves are no doubt due to the percepti-

[2] The limens were calculated by means of Newhall's interpolation formula Lr_4. Newhall, S. M., Linear interpolation vs. the constant process. *Amer. Jour. of Psychol.*, 1927, **38**, pp. 390–402.

[3] Oberly, H. Sherman, The range for visual attention, cognition and apprehension. *A. J. Psychol.*, 1924, **35**, 332–352.

bility of the numbers under tachistoscopic conditions. It is interesting to note that the difference of perceptibility as between size 4 mm. and 6.5 mm is much greater than the difference between 6.5 mm and 9.5 mm.

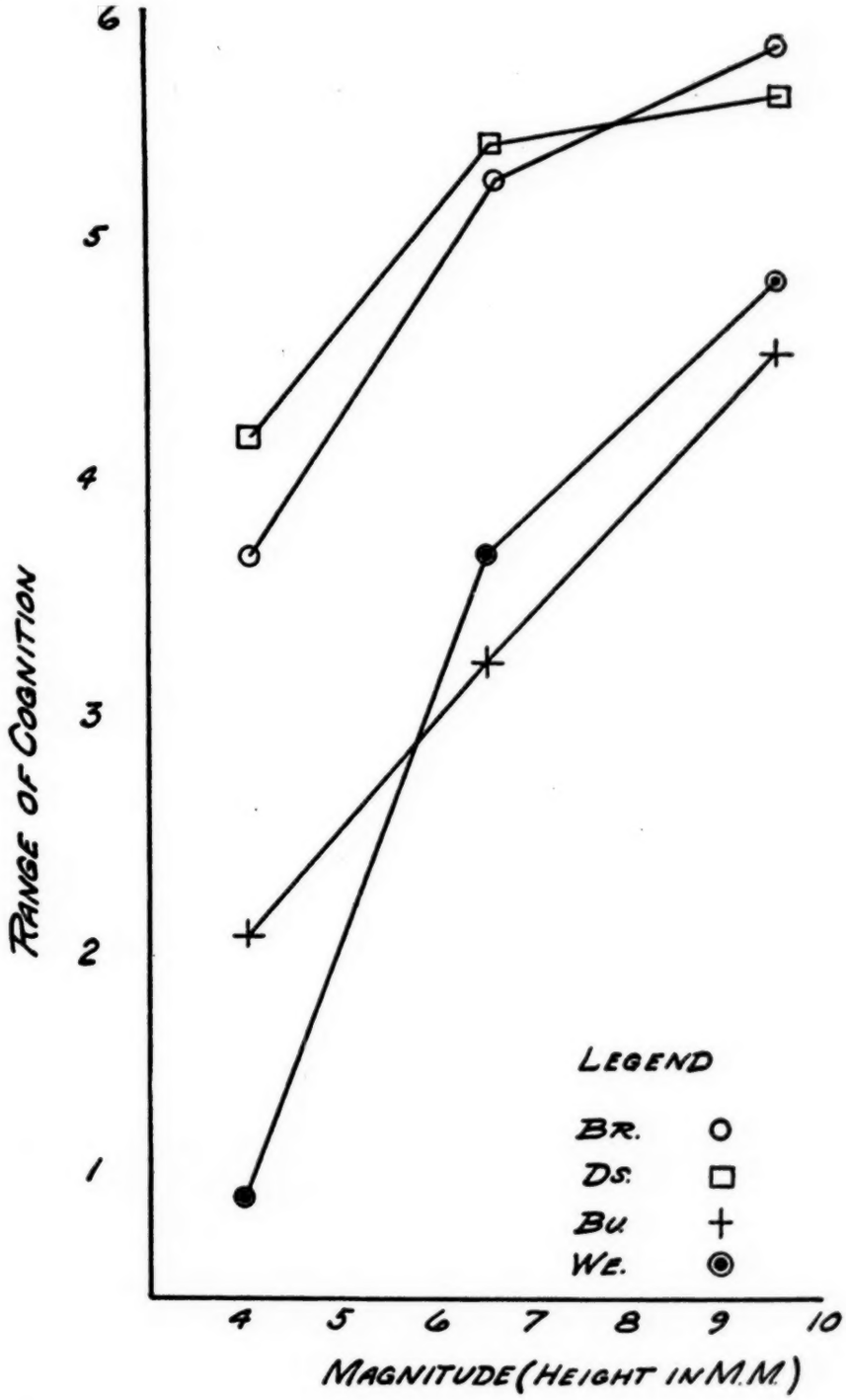


FIG. 1.

The results of the memory experiments should be similar to these curves if perceptibility were a factor. There still remains the possibility of interpreting the memory results on the basis of perceptibility by supposing that the lesser perceptibility of the smaller numbers evoked greater effort and attention and thus by this means the defect in perceptibility was compensated for. This conception, however, proves to be very unlikely in consideration of other available evidence.

An experiment on the memory value of size when only two sizes, 3 and 6 mm., were used indicated that the larger size was better remembered. This experiment was conducted under conditions similar to the memory experiments already reported the only difference being that the method of recognition was used rather than the method of retained members.

The results of this experiment are tabulated in Table IX. Six advanced students in psychology acted as observers. The material consisted of series of 10 two digit numbers of which 5 were 3 mm and 5 were 6 mm in height. One presentation of a series was given and recognition took place immediately after by the method of selecting the 10 numbers seen from a set of 20 which contained the 10 presented numbers.[4]

TABLE IX

NUMBER OF CORRECT RECOGNITIONS

Observer	*L*(6 mm.)	*S*(3 mm.)
Fk	162	162
Mck	146	119
Mur	153	112
Mc	143	116
Mar	157	140
Mil	171	138
	932	787

The results of Table IX may be compared to those of Table II. In both cases six subjects were used and one presentation of the respective memory series were given. The difference was in the manner of measuring memory. Since both the method of retained members and the recogni-

[4] These data are taken from my thesis, *The factor of size for efficiency in memorizing numbers*, presented in partial fulfillment of the degree of Master of Arts at Columbia University, June 1923.

tion method indicate that size has a memory value it is evident that magnitude has a similar result on recognition and recall, and consequently tests of memory employing the method of recognition and those employing retained members are comparable. The only difference would be a matter of degree, since more is recognized than recalled.

To make the comparison of the results of Table II and Table IX quantitatively equivalent I have reduced the results of the latter, which is the result of 3360 items presented, to fit the conditions of 1296, the number of two digit numbers used in Table IX. This is possible without much distortion of fact, because the learning curve under conditions of passive

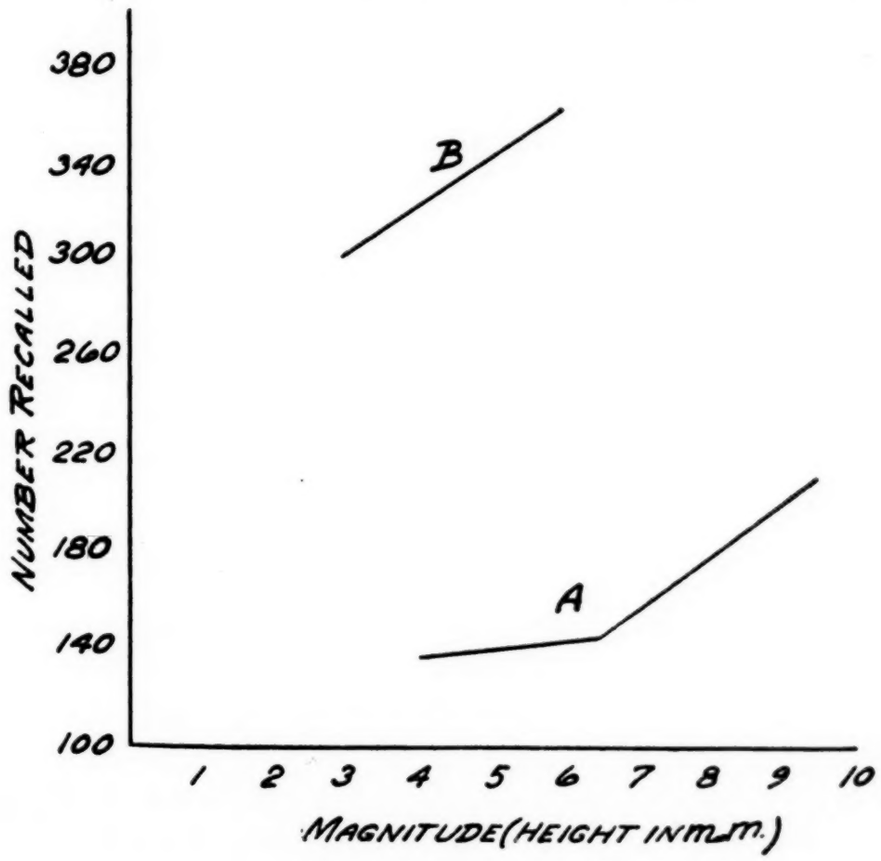


FIG. 2.

memorizing is very gradual. In fact, the practise effect is slight. Taking 1296 as a common measure I find that the results for Table IX should be 362 for large and 302 for small numbers.

In Fig. 2 these results are plotted together with those

of Table II. The abscissæ denotes size of the numbers in terms of height; the ordinates the number of times the numbers were remembered. Curve A shows the results of memorizing numbers varying in three sizes; Curve B, the same, for two sizes.

These curves indicate that magnitudes which are indifferently responded to when three sizes are involved, become differentiated in respect to memory when only the two sizes are used. It may be argued that the indifference shown by the two smaller sizes in the situation of three sizes is really a matter of absolute values, since the sizes were not identical to the magnitudes in the experiment with two sizes. Furthermore it may be insisted that the result may be attributed to the different methods of measuring memory, namely recall and recognition.

Fortunately the literature on memory makes it possible to present examples which indicate that the memory value of size depends upon the organization or the pattern of magnitudes involved. Paula Meyer [5] studied the problem of memory for graphical figures with respect to their position in the retinal field. Incidently the figures used varied in size. Taking all the cases in which figure and position were given correctly it was found that the figures of different size were membered as follows:

5 mm. = 227 10 mm. = 244 15 mm. = 246.

Miss Meyer also made a special study of the factor of size in memory for figures, by using graphical forms of 10 and 40 mm in height. It was found that in case of adults the larger figures were more often recognized.

Floors [6] performed experiments after the method of Paula Meyer with the sole purpose of investigating the effect of size on memory. He used the same kind of material and the same memory procedure. The material consisted of graphical figures which were varied in the following sizes: 1.5, 2.33, 2.73,

[5] Meyer, P., Über die Reproduktion eingeprägter Figuren und ihrer raümlichen Stellengen bei Kindern und Erwachsenen. *Zeits. für Psychol.*, 1912, **64**, S. 34–91.

[6] Floors, M., Über den Einfluss der Grossenvariierung bei Gedächtnisleistungen. *Zeits. für Psychol.*, 1922, **90**, S. 302–328.

3.69, 4.98, 6.73, 9.08, and 12.26 cm. The subjects were presented with these figures, one by one. Each series of figures contained various sizes. Twenty-four hours later a similar series but containing material not previously presented was then given and the subjects were asked to indicate which figures they recognized.

The following table is taken from Floors' article. These data are the result of 48 presentations of each size.

TABELLE III

	Zahl der Wieder erkennungen.	Zahl der Treffer und Teiltreffer bei vorgez. Figur.	Zahl der Treffer u. Teiltreffer bei dargeb. Silbe.
1. bei kl. Figur..... (Grösse 1–4)	116	112	125
2. bei gr. Figur..... (Grösse 5–8)	121	124	134
3. bei Grösse 1.....	24	29	28
4. bei Grösse 8.....	32	33	37

The work of Floors on a wide range of magnitudes indicates that when many sizes are involved the memory value for size is slight. Moreover it is evident from the above table that the difference attributed to size is really the difference between the very largest and the very smallest magnitudes. The sizes intermediate between Grösse 8 and Grösse 1 are indifferently responded to. The work of Meyer in comparing the memory value of magnitudes 10 and 40 mm may be regarded as part of the series of magnitudes used by Floors. The fact that Meyer found the larger of these to be better remembered shows that the memory value of size depends upon the number of magnitudes involved. Since the conditions of the experiments of both Floors and Meyer were practically the same, there can be no objection to the conclusion that the memory value of size is a function of the organization of magnitudes. This fact substantiates my experiments which are statistically more reliable.

The researches considered in this article reveal that the concept of the memory value of size is quite complex. It is not the impressiveness of size, per se, but rather its relative impressiveness with respect to all the other sizes in the experi-

mental situation which determines its memory value. A difference of two magnitudes may be significant for memory when only these two sizes are considered; with an addition of a third size it is possible to reduce the formerly effective size to indifference. Increasing the number of magnitudes tends to wash out the effectiveness of size altogether, and what greater value may be attributed to magnitude is largely attributable to the largest size in the constellation. Hence the memory value of size is an essentially relativistic term, relative to the kind and number of magnitudes involved.

The facts of these experiments are born out by every day life. The tallest building, the highest mountain peak make a stronger impression upon us than the other objects in the field of view. This increased impressiveness is due to factors associated with attention. The largest object is more novel, more interesting, striking against an indifferent background. In the experiments reported it was attempted to hold the attention constant but it is evident that the results obtained are due to the conditions of impression rather than modification of memory traces with the lapse of time because in my work recall was made immediately after presentation of the material. Furthermore, the fact that one size may be effective in one situation and not in another indicates that the impressiveness of size is not of the nature of intensity of stimulation per unit area, and that these units may be added to give greater units. If this were so then size would have a memory value mathematically expressible in the form of an equation indicating a constant relation between memory value and size.

The simplest explanation of these facts is, then, not in terms of size, per se, but rather in terms of attention. Some years ago Peterson [7] showed that the same material could be remembered better when varied in type. He explained this result on the conception that variation in typography introduced interest and novelty into the memory material and thus attention was enhanced. The same sort of thing is involved

[7] Peterson, H. A., On the Influence of Complexity and Dissimilarity on Memory. *Psych. Rev. Mon. Suppl.*, 1909, XII, No. 2.

in these experiments on the memory value of size. To the extent that any size is sufficiently different it becomes more novel, more interesting and more or less separated from the background of the other elements. In this way the memory material becomes organized and a memory pattern or Gestalt is formed.

This explanation may be extended to other facts. It is known that in the work on advertising it has been found and verified that the memory value of relative size corresponds to the size of the advertisements involved. This law holds even when many sizes are involved. These facts have been explained in terms of attention. The situation is such that attention may be directed to one object at the exclusion of the others. There is a tendency for the larger objects to attract and retain the attention; this fact alone is sufficient to explain the results on memory. In the experimental situations here reported the conditions of attention were more constricted. The material was presented successively for the same period of time. The attention was continually focused on the one object in the field of view, thus there was little free play of the attention, and the degree of attention tended to be more or less equalized for all the objects exposed. However, even under these conditions the elements of novelty and interest may enter in to give a certain limited degree of freedom to the distribution of attention. Thus the largest size in the series of magnitudes stands out and is more interesting than the other sizes—or in other words it gets more attention.

Size has a memory value in proportion as it has attention value. On this assumption the changes of the memory value of a definite size as a result of the addition of other magnitudes is readily explained, since such additions merely extend the range of attention indifference or monotony. This attention indifference is essentially greater in the case of stimuli presented successively than it is in the case of simultaneous presentation, and in this fact lies the explanation of memory values for size obtained by the methods of advertising experiments which do not constrain attention, and the laboratory memory experiments which limit its distribution. The

differences in the memory value for absolute and relative size is also founded on this fact since materials of the same size are more monotonous than those involving size differences. Hence any condition that produces monotony or attention indifference tends to lessen the memory value of size.

(Manuscript received October 15, 1930)

THE APTITUDE HYPOTHESIS IN MOTOR SKILLS

BY SIGFRID SEASHORE[1]
University of Oregon

According to present day theories of the nature of individual differences in motor skills it should be possible to develop a battery of motor tests which should indicate likelihood of profiting by training in a complex practical skill. In fact this belief is so strongly held that failures of experimental prediction have usually been attributed to faulty techniques such as unreliabilities of the measures used. In the case of the practical studies this explanation is probably justifiable, especially since in most studies the reliabilities are not determined, or at least are not published. However we cannot neglect the laboratory studies of known reliabilities, (*r* from .80 to .90), and their definite conclusions.

Simple reaction time does not correlate with speed of serial reaction even when the same type of moving visual stimuli and right hand response is used for both, (Farnsworth Seashore Tinker (**3**), Beck (**1**), and Hansen (**4**). Speed of simple reaction does not correlate with speed of tapping a telegraph key for short periods, (Perrin (**5**), Reymert (**7**), Ream (**6**), and Seashore (**8**). These studies suggest the fact that no correlation need be found between two qualitatively similar performances involving the same sense fields and musculatures. However a recent study by the writer (**10**) on the simple reaction time of different musculatures to an auditory stimulus, and a related study by Campbell (**3**) on individual differences in the speed of serial reaction of large muscle groups suggest that relationships involving the same problem and sense fields for different musculatures may be very close (*r* of .90).

[1] The writer wishes to express his appreciation to Mr. Carl Vreeland, Personnel Manager of the Jantzen Knitting Mills at Portland, Oregon, through whose coöperation this study was made possible.

Indications from previous studies then point toward such a high degree of specificity of skills that the prediction of practical skills from a representative test battery would not be feasible. However several points suggest that positive results might be secured from such a study; (1) The common theory of aptitudes as determinants of complex motor skills; (2) The use of a representative test battery would sample a variety of types of coördinations involving the same sense fields and musculatures as the criterion, so that any one of the separate tests might be found to be related to the criterion; (3) The reliability and objectivity of the battery and criterion to be used in the present study were high enough to avoid the criticism used against previous practical studies; (4) An entirely different age range and experimental situation were to be found in the factory as compared with selected university

TABLE I

INTERCORRELATIONS OF MOTOR SKILLS TESTS

	KPR	MR	SD	ST	SP	SR
Koerth Pursuit Rotor						
Motor Rhythm	.295					
Serial Discrimeter	.17	.399				
Speed of Tapping	.014	.252	−.033			
Spool Packer	.03	.03	.287	.354		
Speed Rotor	.161	.134	.236	.477	.295	

CORRESPONDING VALUES OF r AND P.E. WHEN $N = 50$

	P.E.	r	P.E.	r	P.E.
0	.095	40	.08	80	.034
10	.094	50	.072	85	.026
20	.092	60	.061	90	.018
30	.087	70	.049	95	.009
		75	.042		

Test	$r_{\frac{1}{2}-\frac{1}{2}}$	Range	S.D.	Mean
KPR	.932	1003–3146	497.5	2249.5
MR	.783	272–1348	230.2	912.5
SD	.957	545–1530	233.5	1085.5
ST	.866	180– 315	27.2	222.7
SP	.810	470– 670	45	1576.7
SR	.804	1477–2078	151.2	1761.9
Average	.859			

groups; and (5) The skill to be predicted is of a far more complex nature than those used in the laboratory studies.

The Stanford Motor Skills unit which was selected for the testing has several advantages over previous test batteries; (1) the tests are all serial in nature and range in reliability from .78 to .96 tending to fall around .86; (2) the intercorrelations between the separate tests of the unit are low, pointing to the fact that different abilities are being tested (correlations range from −.03 to +.48, clustering around +.20); (3) each test is constructed so as to measure only one type of variable and a unified scoring system is used throughout; (4) in five of the six tests the scoring is automatic and all are objective; (5) the apparatus is standard so that scores from one study may be compared with those from others. The tests which make up the unit are: (1) the Koerth pursuit rotor, accuracy in following a small target revolved by a phonograph motor in a circular path at a high rate of speed; (2) motor rhythm synchrometer, temporal precision in following a simple repeated auditory rhythm on a telegraph key; (3) speed of tapping on a telegraph key; (4) the serial discrimeter, speed in making discriminative finger reactions to number signals appearing as fast as reacted to correctly; (5) the Miles speed rotor, speed of rotary arm, wrist, and finger movements in turniEg a small hand drill having a Veeder counter; and (6) the Brown spool packer, speed of bimanual coördination resembling an industrial operation. The tests were given according to the standard directions from the original article (**9**), with the exception that two experimenters worked together, each giving three of the tests. This permitted the use of two observers at the same time which added to the confidence of the workers and introduced an extra element of competition. Fifty male observers were used, ranging from 16 to 26 years in age with an average of 18 years.

The tests were conducted in the Jantzen Knitting Mills at Portland, Oregon, which is an ideal factory for an experiment of this sort. The morale of the plant is excellent as only the higher class of factory workers are employed. The building is of the one story daylight type, and the very

best of machinery is used throughout. The machinery is run at a high enough speed that a man's capacity for work is never limited by the capacity of his machine. Pay is entirely on the basis of piece rates, worked out on the Bedaux system. The amount of work done by each man is posted daily, appearing in red whenever he drops below his standard. There is a strong spirit of rivalry, but each man takes an interest and pride in the accomplishments of the others. In general there is a feeling of an organization rather than a mere group of employees. All men hired by the plant are started in the winding room and are promoted to other more technical processes only on the basis of amount of work done. Since operation of the winding machines is the starting point in the plant, success in the operation of a winding machine was taken as the practical complex skill to be predicted.

Winding is a very high speed process. Yarn fed into the machine from spindles, is automatically tested for strength and proper thickness, coated with parafine, and rewound on large spools. The operator starts by putting the spindle on a vertical rod about six inches from the floor, and working up to a four foot level, threads the machine by looping the yarn around two hooks, and between two rolls of paraffine. Then, after starting the yarn on the winding spool, he releases a hand clutch which sets the machine in motion. Breakage of the yarn automatically stops the machine. This time the operator picks up the yarn at about a three foot level, and after rethreading the machine takes the loose end of yarn from the winding spool and ties a knot. The machine is then thrown back into gear by the release of the clutch. A beginning operator handles only one machine, but soon becomes proficient enough to handle about six, and an especially fast man can handle eight machines. Since the yarn in each machine breaks on an average of about once every forty seconds, while a spindle lasts for about ten minutes, the main work consists in the tying of the knots, and threading the upper part of the machine. The "weavers" knot which is used is very complicated so that a great deal of difficulty is usually experienced in learning to use it. The knot must be

just as strong as any other part of the yarn, and in addition must be small enough to pass through the needles of the knitting machines. New operators are always required to spend an entire day practicing the knot before they are allowed to start on the machine.

An entirely objective criterion of success in operating the winding machines was secured by means of the piece rate which in this operation is measured by weight. The rating used was the average amount of work done per hour during the last two week period of the first three months work in the plant for each individual. Scores ranged from 25 to 76 Bedaux units. The reliability of this rating taken by correlating the average for this period with the average for the two weeks period just preceding was found to be .82 ± .03. Use of this criterion eliminated all subjective elements such as ratings by foremen, which have been used in similar studies.

TABLE II

CORRELATIONS BETWEEN TESTS AND CRITERION

KPR	.072
MR	.0
SD	.123
ST	.216
SP	.301
SR	.362
Battery	.138

As shown in Table II correlations of the separate tests with the criterion ranged from −.07 for the Koerth pursuit rotor to +.36 for rotary speed with an average of +.16. The correlation between the unit as a whole and the criterion was +.14 ± .10. Considering the high reliabilities of both the tests and the criterion this non-significant correlation may be offered as definite evidence against the hypothesis of a group of basic motor skills as determinants of practical machine skills, and as pointing toward a theory of highly specific motor skills. It is significant to note that the relationship is so low that even the application of Thurstone's methods of critical levels offered no prediction. That is, no matter how low the score found on any one of the tests, it was impossible to say that the observer could not succeed in the

machine operation. As further evidence on this point test scores were averaged for ten men who had been dropped from the plant for failure to reach a satisfactory standard. A comparison of the means of this group with those of the successful operators showed that each group excelled in three of the tests, but in no case was the difference between the means over .10 sigma.

As a minor part of the study, time tests were taken of the actual winding machine operation in order to determine the validity of short time tests at maximum speed in predicting average performance in the same operation over a long period of time. To test this out the five best and the five poorest machine operators were timed and compared. Each man was timed separately using (1) ten trials of tying five knots each, and (2) ten trials of tying five knots and starting the machine. Comparison of the means of the two groups showed that the poorer group excelled the ordinarily fast group on these short time tests. Differences between the means were not significant, but the fact that the ordinarily fast group did not excell on these tests brings the theory of short time sampling very seriously under question.

To explain these peculiar results the hypothesis was proposed that the difference over a long period of time might be caused by fatigue on the part of the poorer workers due to excess or wrong movements. An attempt was made to bring out these differences if they existed, by means of slow motion camera studies. In this study pictures were taken of two good men and two poor men to determine, if possible, any differences in types of movements used. However no such differences were apparent on the short time tests. It is quite possible that this is due to faulty techniques in the interpretation of pictorial results.

Another possibility to be considered is that of pace setting: That is, individual differences in the practical skills might be brought about by the fact that each man tends to set a certain pace in relation to the rest of the group. This hypothesis could be tested out by correlating the ratings of the same operators on two different machine processes. Further interpretation must await a study of this type.

REFERENCES

1. BECK, F. L., *The relationships existing among simple action, serial action, and speed in reading.* (Univ. of Ore. study not yet published, 1930.)
2. CAMPBELL, M., *Individual differences in the speed of serial reaction of large muscle groups.* (Thesis not yet published, Univ. of Ore. Library, 1930.)
3. FARNSWORTH, P. R., Seashore, R. H., and Tinker, M. A., Speed in simple and serial action as related to performance in certain "intelligence" tests. *J. Genet. Psychol. and Ped. Sem.*, 1927, **34**, 537–551.
4. HANSEN, C. E., Serial action as abasic measure of motor capacity. *Psychol. Monog.*, 1922, **31**, 320–382.
5. PERRIN, F. A. C., An experimental study of motor ability. *J. Exper. Psychol.*, 1921, **4**, 24–57.
6. REAM, M. J., The tapping test: a measure of motility. *Psychol. Monog.*, 1922, **31**, 293–319.
7. REYMERT, M. L., The personal equation in motor capacities. *Scand. Scient. Rev.*, 1923, **2**, 177–194.
8. SEASHORE, R. H., Individual differences in motor skills. *J. General Psychol.*, 1930, **3**, 38–66.
9. SEASHORE, R. H., Stanford motor skills unit. *Psychol. Monog.*, 1928, **39**, 51–66.
10. SEASHORE, R. H., Seashore, S. H., Stephenson, D., Walker, R. W., and Wickersham, R., *The "personal equation" in the field of motor performances.* (Study not yet published, 1930, Univ. of Ore.)

(Manuscript received September 8, 1930)

GROUP JUDGMENTS IN THE FIELD OF PERSONALITY TRAITS

BY MAPHEUS SMITH

University of Kansas

The article by Gordon, Group judgments in the field of lifted weights (*Journal of Experimental Psychology*, 1924), induced the writer to treat certain judgment data by similar methods in order to learn if the principles formerly discovered operate in judgments of personality traits. The material judged consisted of records of the behavior of a child of seven years, these records being detailed reports by a trained observer of activities of the child in play situations of a child-placing home. The complete case record consisted of twenty separate reports, each of at least five minutes duration, together with a 'note' giving in more detail the situation at the moment. The observation record proper was inscribed as the activity occurred, the note being added immediately at the close of the observation report. Three different observers contributed to the observation records of the child in question, the observations being made on various days at varying hours.

The judgments were made by 50 first and second year university students of psychology and sociology, and, for the most part, represent the reactions of academic groups in the experiment. The case material was given out in typewritten form with the directions to read the material carefully one time, noting the occurrence of behavior and personality traits such as were included in a prepared list of traits. After the case was read the list of behavior types and personality traits was checked for the degree of occurrence of each trait the judge remembered. The four degrees listed were 'great,' 'medium,' 'small,' and 'none.' Sixty different behavior and personality traits made up the list, including such activities as locomotion, manipulation of objects, looks at others, talks with others, alone, with others, touches others; and such personality traits as laziness, intelligence, leadership,

friendliness, fear, suggestibility, cheerfulness, and seeks recognition. The judges were urged to only rate such traits as were particularly noted. No explanation of the meanings of names of traits was given.

The statistical treatment consisted of numerical counts of the occurrence of behavior types and of personality traits in the observations and notes. The frequency of occurrence was then translated into percentages for each group of traits mutually exclusive and, among them, totalling 100 per cent For example, 'alone' plus 'with children' plus 'with adults' equal 100 per cent, while all activities taken together ('locomotion,' 'manipulation, 'touching others,' and so on) equal 100 per cent. This percentage distribution for the sixty traits was then correlated with each individual judgment, ten mutually exclusive groups of five individuals, five mutually exclusive groups of ten individuals, two mutually exclusive groups of twenty individuals, and the entire group of fifty individuals, the formula used being that for obtaining the Pearson coefficient *r*. This general grouping procedure was in agreement with that of Gordon in the paper mentioned above.

Table I gives a summary of the correlations. The range was from .055 to .633 for individuals, .357 to .620 for groups of five, .434 to .589 for groups of ten, while the two groups of twenty were .458 and .528 respectively. The conclusion that is most evident is that the amount of increase by groups recorded by Gordon was not discovered in the present study. The one group of fifty cases does not approach a perfect correlation nor is it as high as the highest individual correlation, both of which facts the previous investigation disclosed. Further, as indicated by the small increase of the *r* for the fifty judges as against the mean *r* for twenty judges—.493 to .510—there is small hope of a much larger increase in the correlation with a large increase in the number of cases.

However, examination does show that the grouped correlations are distinctly higher than the average of the scores composing the group. For example, only six individual scores surpass the correlation for the entire group of fifty judgments.

And again, the mean of the means of the individual *r*s for groups of five, ten, twenty, and fifty judgments are all the same, .366, while the mean *r* by grouped scores of five, ten, twenty, and fifty judgments are, respectively, .461, .489, .493, and .510. (Cf. Table I.) In short, the data presented indi-

TABLE I

Intervals	Frequency				
	Single Cases	Groups of 5	Groups of 10	Groups of 20	Groups of 50
1.00–.96					
.95–.91					
.90–.86					
.85–.81					
.80–.76					
.75–.71					
.70–.66					
.65–.61	1	1			
.60–.56	1		1		
.55–.51	4	2		1	1
.50–.46	6	1	3	1	
.45–.41	9	4	1		
.40–.36	3	2			
.35–.31	12				
.30–.26	8				
.25–.21	1				
.20–.16	2				
.15–.11	2				
.10–.06	1				
.05–.01					
Totals	50	10	5	2	1
Mean *r*	.366	.461	.489	.493	.510

cate that similar factors functioned in the study of personality trait judgments as were operative in the earlier study, but occur in considerably less degree in the present instance.

Search for an explanation of an increase in group correlations which tends apparently to approach the highest individual correlations as a limit leads most logically to a comparison of mean *r*s for five judgments and *r*s of groups of five, since the same general principles would be supposed to operate in the larger groupings. We find in each of the ten possible comparisons that of the terms in the formula $r = \Sigma xy/n\sigma_x\sigma_y$ the σx remains constant since the x variable was chosen to represent the statistical count of the case records, but that the y^2 figure decreases in every instance by a mean of

$-21{,}633.06$ (range $-14{,}854.07$ to $-30{,}135.82$), causing σy to decrease. Reference to the other variable, Σxy, indicates that it also decreases but not so much nor so consistently, the mean decrease being -475.40 (range $+1094.14$ to -2349.96) The greater decrease of the σy accounts for the increase in correlations, as the mean rs become group rs. The decrease of the Σxy prevents the correlation from rising to a degree more comparable with Gordon's results. Beyond this explanation the present data offer no suggestion.

The question as to why the highest individual r and the highest group r are much lower than indicated in Gordon's work cannot be definitely answered as yet. The two most plausible reasons are, first, a difference of definitions of traits by the judges and the investigator, and second, the use of too long a list of traits for careful and uniform work on the entire list. When these difficulties are eliminated the r of both the highest individual score and of the largest grouping of scores may be expected to approach a more perfect correlation, although the potency of the grouping of scores for increasing correlations is certainly indicated by the personality trait judgment method herein described.

(Manuscript received September 1, 1930)